CONSTRAINED BAYES AND EMPIRICAL BAYES ESTIMATORS UNDER SQUARED ERROR AND BALANCED LOSS FUNCTIONS

By

MYUNG JOON KIM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

Dedicated to the God for giving me endless love,

my parents for giving me endless support,

my brother for giving me endless advice,

my wife for giving me endless encouragement,

my daughter for giving me endless smile,

with all my love from Jesus.

ACKNOWLEDGMENTS

This dissertation could not have been written without Distinguished Professor, Malay Ghosh, who not only served as my Ph.D. committee chair, but also encouraged and challenged me throughout my graduate program. I thank him sincerely.

The other committee members, Professors Ronald Randles, Andrew Rosalsky, Cynthia Garvan and Beverly Brechner, patiently guided me through the dissertation process, never accepting less than my best efforts. I thank them all also.

I also would like to give my special thanks to the Department of Statistics, which supported me financially and offered me a great knowledge in statistics throughout my years at the University of Florida.

I express heartfelt gratitude to my father, Chang-Sun Kim, my mother, Keum-Ja Kim and my brother, Myung-Suk Kim, for their never-ending support, advice and love throughout my studies. I will never forget their sincere minds towards me.

Most importantly, I thank my wife, Su-Yeon Hwang who showed me her unwavering love and patience. Her watchful eyes enabled me to continue my study and finally finish my dissertation.

I also would like to express my enormous love to my lovely daughter, Min-Ji, who made my study meaningful.

Finally and especially I thank my and our God!

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

CONSTRAINED BAYES AND EMPIRICAL BAYES ESTIMATORS UNDER SQUARED ERROR AND BALANCED LOSS FUNCTIONS

By

Myung Joon Kim

May 2004

Chair: Malay Ghosh
Major Department: Statistics

Bayesian and empirical Bayesian methods have become quite popular in the theory and practice of statistics in the last two decades. In particular, hierarchical and empirical Bayesian methods are very suitable in the context of simultaneous estimation when there is a genuine need for "borrowing strength."

Often, however, the goal is not just to produce an ensemble of estimates simultaneously for several parameters but also to produce a set of estimates whose empirical histogram estimates well the histogram of population parameters. However, in a general framework, the histogram of the posterior means of coordinate-specific parameters is underdispersed as an estimate of the histogram of parameters. This requires adjustment of Bayes and empirical Bayes estimators. One way to meet the twin objectives as mentioned earlier is to match the first two empirical moments of the Bayes estimates with the posterior means of the mean and variance of the population parameters. The resulting estimators are referred to as constrained Bayes estimators.

It is also known that least squares estimators reflect goodness of fit consideration, while quadratic losses reflect solely precision of estimation. Thus there

is a need to provide a framework within the tradeoff between goodness of fit and precision of estimation. With a consideration of this need, balanced loss functions are introduced which reflect two criteria–goodness of fit and precision of estimation.

This dissertation focuses on constrained Bayes and empirical Bayes estimators with asymptotic measures of precision associated with these estimators. We consider both the squared error loss and the balanced loss. Estimators are derived under several situations, such as the one-parameter exponential family with conjugate priors, with particular emphasis on the normal-normal case, and the balanced random effects normal ANOVA model. Also asymptotic measures of precision associated these estimators are derived which are valid up to a specified order of approximation.

CHAPTER 1
INTRODUCTION

One of the main objectives of this dissertation is to obtain the asymptotic mean squared errors (MSE's) of constrained Bayes and empirical Bayes estimators which are correct up to a certain order, and obtain estimates of these MSE's which are asymptotically unbiased. In addition, constrained Bayes and empirical Bayes estimators are found under balanced loss functions. Their asymptotic Bayes risks are calculated, and asymptotically unbiased estimators of these Bayes risks are obtained.

1.1 Literature Review

Bayesian techniques are widely used for simultaneous estimation of several parameters in compound decision problems. A well-known example is small area estimation where interest lies in simultaneous estimation of means or other parameters of interest, say, for counties, census tracks or other local areas. Under any quadratic loss, the Bayes estimates turn out to be the posterior means of the parameters of interest.

Often, however, the objective is not only to produce an ensemble of parameter estimates under a certain loss, but also to ensure that the histogram of the estimates is somewhat close to the histogram of the population parameters. For example, if $X_i|\theta_i$ are independent $N(\theta_i, 1)$ and θ_i are *iid* $N(\mu, A)$, $(i = 1, 2, ..., m)$, then assuming squared error loss, the Bayes estimator of $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)^T$ is the posterior mean $((1-B)X_1 + B\mu, ...(1-B)X_m + B\mu)^T = (1-B)\boldsymbol{X} + B\mu\mathbf{1}_m$, where $B = (1+A)^{-1}$, $\boldsymbol{X} = (X_1, ..., X_m)^T$ and $\mathbf{1}_m$ is an m-component column vector with each element equal to 1. On the other hand, the optimal Bayes estimator of the population histogram of parameters, namely, $m^{-1}\sum_{i=1}^m I_{[\theta_i \le t]}$

(I being the usual indicator function) is given by $m^{-1}\sum_{i=1}^{m} P(\theta_i \leq t|X_i) = m^{-1}\sum_{i=1}^{m} \Phi[\{t-((1-B)X_i - B\mu)\}/(1-B)^{1/2}]$ where Φ is the distribution function of the $N(0,1)$ variable, since $\theta_i|X_i$ are independent $N((1-B)X_i + B\mu, 1-B)$. As pointed out by Louis (1984), who examined an example from hypertension detection and follow-up study, this is the situation in subgroup analysis, where the problem is not only to estimate the different components of a parameter vector, but also to identify the parameters above and other parameters below a specified cutoff point. More generally, one may be interested in classifying the parameters into several categories. Spjotvell and Thomsen (1987) documented that it is possible to adopt a modified Bayesian procedure to improve on the posterior means as estimates of proportions.

The twin objectives mentioned in the previous paragraph are usually conflicting in the sense that one is often achieved at the expense of the other.

The posterior means of the parameters of interest are the optimal estimates under any quadratic loss. However, it can be shown in a general framework that the histogram of the posterior means of coordinate-specific parameters is underdispersed as an estimate of the histogram of parameters. Accordingly, the histogram of posterior means is clearly inappropriate to estimate the parameter histogram. Indeed, no single set of estimates can simultaneously optimize the two goals as mentioned in the previous paragraph. However, in many policy settings, communication and credibility are enhanced by reporting a single set of estimates with good performance for both the goals. Thus, there is a need to find a set of estimates which is suboptimal according to each one of the two criteria, but serves as a very useful compromise between the two.

To this end, Louis (1984) proposed a constrained Bayes method which matches the first two empirical moments of the Bayes estimates of the normal means with the corresponding moments derived from the posterior histogram, and minimizes

the squared distance of the parameters and estimates subject to these constraints. Ghosh (1992) generalized Louis's findings to obtain results for any arbitrary distribution, not necessarily normal. The resulting estimates are referred to as constrained Bayes (CB) estimators.

Ghosh (1992) derived such estimators for the one-parameter natural exponential family of distributions with quadratic variance functions (NEF-QVF) when the parameters of interest were the population means. Also, empirical Bayes (EB) analogues of the CB estimators, referred to as CEB estimators, were developed in the normal case, and these estimators were analogues of the celebrated James-Stein estimators.

Consider the situation where the data are denoted by $\boldsymbol{x}$ and $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)^T$ is the parameter of interest. Let $\boldsymbol{e}^B(\boldsymbol{x}) = (e_1^B(\boldsymbol{x}), ..., e_m^B(\boldsymbol{x}))^T$ denote the Bayes estimate of $\boldsymbol{\theta}$ under any quadratic loss based on data $\boldsymbol{x}$. Our objective is to find the Constrained Bayes (CB) estimate $\boldsymbol{e}^{CB}(\boldsymbol{x}) = (e_1^{CB}(\boldsymbol{x}), ..., e_m^{CB}(\boldsymbol{x}))^T$ of $\boldsymbol{\theta}$, where $\boldsymbol{e}^{CB}(\boldsymbol{x})$ minimizes

$$E[\sum_{i=1}^{m}(\theta_i - t_i)^2|\boldsymbol{x}] \tag{1.1}$$

within the class of all estimates $\boldsymbol{t}(\boldsymbol{x}) = \boldsymbol{t} = (t_1, ...t_m)^T$ of $\boldsymbol{\theta}$ that satisfy

$$(a)\ E(\bar{\theta}|\boldsymbol{x}) = m^{-1}\sum_{i=1}^{m} t_i(\boldsymbol{x}) = \bar{t}(\boldsymbol{x}), \quad (say) \tag{1.2}$$

$$(b)\ E[\sum_{i=1}^{m}(\theta_i - \bar{\theta})^2|\boldsymbol{x}] = \sum_{i=1}^{m}[t_i(\boldsymbol{x}) - \bar{t}(\boldsymbol{x})]^2. \tag{1.3}$$

It is easy to see that the estimate $\boldsymbol{e}^B(\boldsymbol{x})$ of $\boldsymbol{\theta}$ satisfies (1.2) . However, it does not satisfy (1.3). To see this we define $\boldsymbol{I}_m$ as the identity matrix of order m, $\mathbf{1}_m$ as the m-component column vector with each element equal to 1, $\boldsymbol{J}_m = \mathbf{1}_m\mathbf{1}_m^T$, and calculate

$$\begin{aligned}
E[\sum_{i=1}^{m}(\theta_i - \bar{\theta})^2 | \boldsymbol{X}] &= tr[(\boldsymbol{I}_m - \frac{1}{m}\boldsymbol{J}_m)E(\boldsymbol{\theta}\boldsymbol{\theta}^T|\boldsymbol{X})] \\
&= tr[(\boldsymbol{I}_m - \frac{1}{m}\boldsymbol{J}_m)\{V(\boldsymbol{\theta}|\boldsymbol{X}) + E(\boldsymbol{\theta}|\boldsymbol{X})E(\boldsymbol{\theta}|\boldsymbol{X})^T]\} \\
&= tr[(\boldsymbol{I}_m - \frac{1}{m}\boldsymbol{J}_m)V(\boldsymbol{\theta}|\boldsymbol{X})] + \sum_{i=1}^{m}(e_i^B(\boldsymbol{X}) - \bar{e}^B(\boldsymbol{X}))^2 \\
&= tr[V(\boldsymbol{\theta} - \bar{\theta}\boldsymbol{1}_m|\boldsymbol{X})] + \sum_{i=1}^{m}(e_i^B(\boldsymbol{X}) - \bar{e}^B(\boldsymbol{X}))^2 \\
&> \sum_{i=1}^{m}(e_i^B(\boldsymbol{X}) - \bar{e}^B(\boldsymbol{X}))^2.
\end{aligned}$$

The above points out very clearly the limitations of usual Bayes estimates in estimating the true variation among the θ_i's.

However, it is possible to find a vector of estimates $\boldsymbol{\theta}$ which minimizes (1.1) subject to (1.2) and (1.3). A theorem to this effect is proved in Ghosh (1992). For stating this theorem, we need a few notations. Let

$$H_1(\boldsymbol{x}) = tr[V(\boldsymbol{\theta} - \bar{\boldsymbol{\theta}}1_m|\boldsymbol{x})] = tr[(\boldsymbol{I}_m - m^{-1}\boldsymbol{J}_m)V(\boldsymbol{\theta}|\boldsymbol{x})]$$

$$H_2(\boldsymbol{x}) = \sum_{i=1}^{m}(e_i^B(\boldsymbol{x}) - \bar{e}^B(\boldsymbol{x}))^2.$$

The main result of Ghosh (1992) is now stated as follows.

Let $\mathcal{X}_0 = \{\boldsymbol{x} : H_2(\boldsymbol{x}) > 0\}$ then for $\boldsymbol{x} \in \mathcal{X}_0$, the solution $\boldsymbol{t}$ of (1.1) subject to (1.2) and (1.3) is given by $\boldsymbol{e}^{CB}(\boldsymbol{x}) = (e_1^{CB}(\boldsymbol{x}), ...e_m^{CB}(\boldsymbol{x}))^T$, where

$$e_i^{CB}(\boldsymbol{x}) = ae_i^B(\boldsymbol{x}) + (1-a)\bar{e}^B(\boldsymbol{x}), \quad i = 1, ..., m,$$
$$a \equiv a(\boldsymbol{x}) = [1 + H_1(\boldsymbol{x})/H_2(\boldsymbol{x})]^{1/2}. \tag{1.4}$$

We shall refer to $\boldsymbol{e}^{CB}$ as the Constrained Bayes (CB) estimate of $\boldsymbol{\theta}$. Equation (1.4) has the deceptive appearance of expressing the components of $\boldsymbol{e}^{CB}$ as convex combinations of the Bayes estimates e_i^B's and their average. This is not so, because

a exceeds 1. Also, in many situations–especially in discrete cases–there is a positive probability that $H_2(\boldsymbol{x})$ is 0; that is, $e_1(\boldsymbol{x}) = \cdots = e_m(\boldsymbol{x})$. Although $\boldsymbol{e}^{CB}$ remains undefined with positive probability in such instances, an asymptotic (as $m \to \infty$) version of such estimators still may be meaningful. Ghosh (1992) showed this for the binomial and Poisson examples. Now we shall discuss CB estimators for the one-parameter exponential family.

Suppose that $X_1, ...X_m$ are m independent random variables, where X_i has the pdf (with respect to some σ-finite measure) given by

$$f_{\phi_i}(x_i) = exp(n\phi_i x_i - n\psi(\phi_i)), \quad i = 1, ..., m.$$

Each X_i can be viewed as the average of m iid random variables, each having a pdf belonging to a one-parameter exponential family. It is assumed that $\psi(\cdot)$ is twice differentiable in its argument. The objective is to estimate $\theta_i = E_{\phi_i}(X_i) = \psi'(\phi_i), i = 1, ..., m$. Assume the independent conjugate priors

$$g(\phi_i) = exp(\nu\phi_i\mu - \nu\psi(\phi_i))$$

for the ϕ_i's. Then under quadratic loss, the Bayes estimates of θ_i's are given by

$$e_i^B(\boldsymbol{x}) = E(\theta_i|\boldsymbol{x}) = E[\psi'(\phi_i)|\boldsymbol{x}] = (1 - B)x_i + B\mu, \tag{1.5}$$

where $B = \nu/(n + \nu)$. Also, from the posterior distribution of θ_i, integration by parts gives

$$V(\theta_i|\boldsymbol{x}) = V[\psi'(\phi_i)|x_i] = (n + \nu)^{-1}E[\psi''(\phi_i)|x_i] = q_i \quad (say). \tag{1.6}$$

It follows from (1.5) that $H_2(\boldsymbol{x}) = (1 - B)^2 \sum_{i=1}^m (x_i - \bar{x})^2$, whereas from (1.6), one gets $H_1(\boldsymbol{x}) = (1 - m^{-1}) \sum_{i=1}^m q_i$. Then the quantity "$a$" is determined from (1.4).

Further simplification in calculating (see Morris, 1982) H_1 is possible when X_i's are generated from QVF (quadratic variance function) subfamily of the

natural exponential family. Then,

$$\psi''(\phi_i) = v_0 + v_1\psi'(\phi_i) + v_2(\psi'(\phi_i))^2 = v_0 + v_1\theta_i + v_2\theta_i^2, \quad 1 \le i \le m, \tag{1.7}$$

where v_0,v_1 and v_2 are not simultaneously 0's and $v_2 < n + \nu$. Then, using (1.6) and (1.7),

$$q_i = (n + \nu - v_2)^{-1}[v_0 + v_1 e_i^B(\boldsymbol{x}) + v_2(e_i^B(\boldsymbol{x}))^2]$$

so that

$$H_1(\boldsymbol{x}) = (m-1)(n+\nu-v_2)^{-1}[v_0 + v_1\bar{e}^B(\boldsymbol{x}) + v_2\{(\bar{e}(\boldsymbol{x}))^2 + m^{-1}H_2(\boldsymbol{x})\}].$$

Consequently, for $\boldsymbol{x} \in \mathcal{X}_0$,

$$\begin{aligned} a^2(\boldsymbol{x}) &= [1 + v_2(n+\nu-v_2)^{-1}(1-m^{-1})] \\ &+ (m-1)(n+\nu-v_2)^{-1}[v_0 + v_1 e_i^B(\boldsymbol{x}) + v_2(e_i^B(\boldsymbol{x}))^2]/H_2(\boldsymbol{x}). \end{aligned}$$

In the normal example $v_0 = \sigma^2$ (known), whereas $v_1 = v_2 = 0$. Then

$$\begin{aligned} a^2(\boldsymbol{x}) &= 1 + (m-1)(n+\nu)^{-1}(1-B)^{-2}/\sum_{i=1}^{m}(x_i - \bar{x})^2 \\ &= 1 + (m-1)n^{-1}\sigma^2/[(1-B)\sum_{i=1}^{m}(x_i - \bar{x})^2]. \end{aligned}$$

In the normal case the probability that all the X_i's are equal is 0.

Ghosh and Kim (2002) have extended the result of Ghosh (1992) when the parameters themselves are vector-valued.

Suppose $\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_m$ are the m vector-valued parameters of interest and $\boldsymbol{e}_1^B(\boldsymbol{x}), ..., \boldsymbol{e}_m^B(\boldsymbol{x})$ are the corresponding Bayes estimates based on the data $\boldsymbol{x}$ under any quadratic loss. Then writing $\bar{\boldsymbol{\theta}} = m^{-1}\sum_{i=1}^{m}\boldsymbol{\theta}_i$, one gets

$$E[\sum_{i=1}^{m}(\boldsymbol{\theta}_i - \bar{\boldsymbol{\theta}})(\boldsymbol{\theta}_i - \bar{\boldsymbol{\theta}})^T|\boldsymbol{x}] = \boldsymbol{H}_1(\boldsymbol{x}) + \boldsymbol{H}_2(\boldsymbol{x}), \tag{1.8}$$

where

$$\boldsymbol{H}_1(\boldsymbol{x}) = \sum_{i=1}^{m} V(\boldsymbol{\theta}_i|\boldsymbol{x}) - mV(\bar{\boldsymbol{\theta}}|\boldsymbol{x}),$$

$$\boldsymbol{H}_2(\boldsymbol{x}) = \sum_{i=1}^{m}[\boldsymbol{e}_i^B(\boldsymbol{x}) - \bar{\boldsymbol{e}}^B(\boldsymbol{x})][\boldsymbol{e}_i^B(\boldsymbol{x}) - \bar{\boldsymbol{e}}^B(\boldsymbol{x})]^T.$$

Thus the posterior mean of the population variability of the $\boldsymbol{\theta}_i$'s given in the left hand side of (1.8) exceeds the corresponding variability among the e_i^B's by $m^{-1}\boldsymbol{H}_1(\boldsymbol{x})$. In the above and in what follows we say that two symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ satisfy the relationship $\boldsymbol{A} \geq \boldsymbol{B}$ if $\boldsymbol{A} - \boldsymbol{B}$ is non-negative definite. We will denote the (i, j)th element of $\boldsymbol{H}_1(\boldsymbol{x})$ by $H_{1ij}(\boldsymbol{x})$ and that of $\boldsymbol{H}_2(\boldsymbol{x})$ by $H_{2ij}(\boldsymbol{x})$.

Generalizing the formulation of Ghosh (1992), the objective is to find $\boldsymbol{t}_1, ..., \boldsymbol{t}_m$ which minimize

$$E[\sum_{i=1}^{m}(\boldsymbol{\theta}_i - \boldsymbol{t}_i)(\boldsymbol{\theta}_i - \boldsymbol{t}_i)^T|\boldsymbol{x}] \tag{1.9}$$

subject to

$$E(\bar{\boldsymbol{\theta}}|\boldsymbol{x}) = m^{-1}\sum_{i=1}^{m}\boldsymbol{t}_i(\boldsymbol{x}) = \bar{\boldsymbol{t}}(\boldsymbol{x}); \tag{1.10}$$

$$E[\sum_{i=1}^{m}(\boldsymbol{\theta}_i - \bar{\boldsymbol{\theta}})(\boldsymbol{\theta}_i - \bar{\boldsymbol{\theta}})^T|\boldsymbol{x}] = \sum_{i=1}^{m}[\boldsymbol{t}_i(\boldsymbol{x}) - \bar{\boldsymbol{t}}(\boldsymbol{x})][\boldsymbol{t}_i(\boldsymbol{x}) - \bar{\boldsymbol{t}}(\boldsymbol{x})]^T. \tag{1.11}$$

As noted already, the usual Bayes estimates $\boldsymbol{e}_1^B(\boldsymbol{x}), ..., \boldsymbol{e}_m^B(\boldsymbol{x})$ of $\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_m$ satisfy (1.10) but not (1.11). The following argument shows how a simple modification of $\boldsymbol{e}_1^B(\boldsymbol{x}), ..., \boldsymbol{e}_m^B(\boldsymbol{x})$ provides the desired solution.

Let $\boldsymbol{e}_1^B(\boldsymbol{x}), ..., \boldsymbol{e}_m^B(\boldsymbol{x})$ denote the Bayes estimates of $\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_m$ under any quadratic loss based on the data $\boldsymbol{x}$. Let $\mathcal{X}_0 = \{\boldsymbol{x} : H_{2jj}(\boldsymbol{x}) > 0 \text{ for all } j = 1, ..., m\}$.

Then for $\boldsymbol{x} \in \mathcal{X}_0$, the solution $\boldsymbol{t}_1, ..., \boldsymbol{t}_m$ of (1.9) subject to (1.10) and (1.11) is given by $\boldsymbol{e}_1^{CB}(\boldsymbol{x}), ..., \boldsymbol{e}_m^{CB}(\boldsymbol{x})$, where

$$\boldsymbol{e}_i^{CB}(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{e}_i^{B}(\boldsymbol{x}) + (\boldsymbol{I} - \boldsymbol{G})\bar{\boldsymbol{e}}^{B}(\boldsymbol{x}), \quad i = 1, ..., m,$$

where $\boldsymbol{G} \equiv \boldsymbol{G}(\boldsymbol{x})$ is a diagonal matrix with jth diagonal element given by

$$[(H_{1jj}(\boldsymbol{x}) + H_{2jj}(\boldsymbol{x}))/H_{2jj}(\boldsymbol{x})]^{\frac{1}{2}}.$$

Ghosh derived constrained Bayes estimators in a number of situations. He also derived constrained empirical Bayes estimators including the constrained empirical Bayes analogue of the celebrated James-Stein estimators.

James-Stein estimators (James and Stein, 1961) have long been popular among statisticians. The theoretical interest in these estimators stems from their minimaxity and other related properties. On the other hand, practitioners have found these estimators quite appealing in the context of simultaneous estimation of parameters when there is a clear need for borrowing strength. While the original James-Stein estimators shrink the multivariate sample mean towards some prior mean, Lindley's (1962) modification of the same shrinks the sample mean towards some grand average of the component sample means. All these estimators have an interesting empirical Bayes (EB) interpretation.

Efron and Morris (1973) showed how James-Stein estimators arise naturally in an empirical Bayes context. We begin with the situation where $\boldsymbol{X}|\boldsymbol{\theta} \sim N(\boldsymbol{\theta}, \boldsymbol{I}_m)$.

Suppose that the prior distribution for θ_i is $N(\mu_i, A)$. Assuming the squared error loss, the Bayes estimator of θ_i is given by

$$\begin{aligned} E(\theta_i|X_i = x_i) &= \frac{x_i + \mu_i/A}{1 + 1/A} = \frac{A}{A+1}x_i + \frac{1}{A+1}\mu_i \\ &= \mu_i + (1 - \frac{1}{A+1})(x_i - \mu_i) \\ &= \mu_i + (1 - B)(x_i - \mu_i), \end{aligned}$$

where $B = (1 + A)^{-1}$.

Suppose that $\boldsymbol{\mu}$ is known, but A is unknown, and it has to be estimated from the data. Note that $\boldsymbol{X} \sim N(\boldsymbol{\mu}, B^{-1}\boldsymbol{I}_m)$. Thus, marginally, $||\boldsymbol{X} - \boldsymbol{\mu}||^2 \sim B^{-1}\chi^2_m$. Hence,

$$E(\frac{m-2}{||\boldsymbol{X} - \boldsymbol{\mu}||^2}) = B, \quad m \geq 3.$$

Substituting this estimator for B, one gets the estimator for $\boldsymbol{\theta}$ as

$$\boldsymbol{\delta}(\boldsymbol{X}) = \boldsymbol{\mu} + (1 - \frac{m-2}{||\boldsymbol{X} - \boldsymbol{\mu}||^2})(\boldsymbol{X} - \boldsymbol{\mu}).$$

Thus, the estimator $\boldsymbol{\delta}$ shrinks $\boldsymbol{X}$ towards an arbitrary point $\boldsymbol{\mu}$.

Suppose now that $\boldsymbol{\mu}^T = (\mu, ..., \mu)$ and μ and A are unknown. Now, marginally $\boldsymbol{X}$ is $N(\mu\boldsymbol{I}_m, B^{-1}\boldsymbol{I}_m)$. In this case $(\bar{X}, \sum(X_i - \bar{X})^2)$ is complete sufficient for (μ, A). Since $\sum(X_i - \bar{X})^2 \sim B^{-1}\chi^2_{m-1}$, the UMVUE for B is given by $\frac{m-3}{\sum(X_i - \bar{X})^2}$ when $m \geq 4$. Also, $\bar{X}$ is the UMVUE for μ. Since, in this case the Bayes estimator of $\boldsymbol{\theta}$ is

$$\begin{aligned} E(\theta_i | X_i = x_i) &= \frac{X_i + \mu/A}{1 + 1/A} = (1 - B)X_i + B\mu \\ &= \mu + (1 - B)(X_i - \mu), \quad i = 1, ..., m. \end{aligned}$$

Substituting the UMVUE estimators of μ and B, it follows that the empirical Bayes estimator of $\boldsymbol{\theta}$ is

$$\boldsymbol{\theta}^{EB}(\boldsymbol{X}) = \bar{X}\mathbf{1}_m + (1 - \frac{m-3}{\sum(X_i - \bar{X})^2})(\boldsymbol{X} - \bar{X}\mathbf{1}_m).$$

It can be shown that $\boldsymbol{\theta}^{EB}(\boldsymbol{X})$ also dominates $\boldsymbol{X}$ for $m \geq 4$. The estimator $\boldsymbol{\theta}^{EB}(\boldsymbol{X})$ shrinks the usual estimator $\boldsymbol{X}$ of $\boldsymbol{\theta}$ towards $\bar{X}$, and was proposed by Lindley (1962).

In this set-up, constrained empirical Bayes estimators are

$$\boldsymbol{\theta}^{CEB}(\boldsymbol{X}) = a_{EB}\boldsymbol{e}^{EB}(\boldsymbol{X}) + (1 - a_{EB})\bar{\boldsymbol{e}}^{EB}(\boldsymbol{X}),$$

where writing $S^2 = \sum_{i=1}^{m}(X_i - \bar{X})^2/(m-1)$, $\hat{B} = (m-3)/S^2$, $a_{EB}^2 \equiv a_{EB}^2(\boldsymbol{x}) = [1 + \frac{1-\hat{B}}{(1-\hat{B})^2 S^2}] = [1 + \frac{1}{(1-\hat{B})S^2}]$.

Constrained estimators were further generalized by Shen and Louis (1998) who proposed the development of "triple-goal" estimates, those producing a histogram that is a good estimate of the parameter histogram, with induced ranks that are good estimates of parameter ranks and with good performance in estimating unit-specific parameters. They showed that a Bayesian procedure, when suitably modified, would meet all the three criteria, and compared them with posterior means and constrained Bayes estimates. Also Shen and Louis (1998) suggested additional study of empirical Bayes estimators and use of a loss function other than the squared-error loss function.

CB or CEB estimators have mostly been derived under squared error loss. One exception is Cressie (1989), who considered weighted squared error loss for obtaining adjusted census counts. However, as pointed out by Louis (2001), there is a need for developing such estimators for the other losses as well. One such loss considered in this dissertation is the so-called balanced loss function.

Balanced loss functions were introduced by Zellner (1988,1992). Such losses are formulated to reflect two criteria–goodness of fit and precision of estimation. As noted by Zellner (1992), least squares estimators reflect goodness of fit consideration, while quadratic losses (which includes the squared error loss) are geared solely towards precision of estimation.

It is well recognized that sole emphasis on the precision of estimation criterion, for example mean squared error, can lead to biased estimators. In some circumstances bias is not important, but in others, it is critical. Thus there is a need to provide a framework within which the tradeoff between goodness of fit, or lack of bias, and precision of estimation can be considered formally. Zellner (1992) suggested a balanced loss function (BLF) which meets this need.

The BLF is defined as

$$L(\boldsymbol{\theta}, \boldsymbol{d}) = w||\boldsymbol{x} - \boldsymbol{d}||^2 + (1-w)||\boldsymbol{\theta} - \boldsymbol{d}||^2$$

where $||\cdot||$ denotes the Euclidean norm and w is weight.

The first term on the right hand side represents goodness of fit while the second represents precision of estimation. Under BLF, the Bayes estimates are not posterior means of parameters of interest any more. We first consider the estimation of a scalar mean and then estimation of a vector mean relative to balanced loss function (BLF).

Let $\boldsymbol{X}^T = (X_1, X_2, ..., X_m)$, the observation vector satisfies

$$\boldsymbol{X} = \theta \mathbf{1}_m + \boldsymbol{u}$$

where θ is the common mean of the X_i's, $\mathbf{1}_m$ is an $m \times 1$ vector with all elements equal to one, and $\boldsymbol{u}$ is an $m \times 1$ error vector. Our problem is to estimate θ, assuming that a posterior density for θ, with some prior informations are available.

A BLF for θ, denoted $L_B(\tilde{\theta}, \theta)$, where $\tilde{\theta}$ is some estimate, is given by

$$L_B(\tilde{\theta}, \theta) = w(\boldsymbol{x} - \tilde{\theta}\mathbf{1}_m)^T(\boldsymbol{x} - \tilde{\theta}\mathbf{1}_m) + (1-w)(\tilde{\theta} - \theta)^2 m, \tag{1.12}$$

with w having a given value in $[0,1]$. The first term in the right hand side of (1.12) represents goodness of fit while the second represents precision of estimation.

We can re-express (1.12) as follows:

$$L_B(\tilde{\theta}, \theta) = w[\hat{\sigma}^2 + (\tilde{\theta} - \bar{x})^2] + (1-w)(\tilde{\theta} - \theta)^2,$$

where $\hat{\sigma}^2 = (\boldsymbol{x} - \bar{x}\mathbf{1})^T(\boldsymbol{x} - \bar{x}\mathbf{1})/m$ and $\bar{x}$ is the sample mean. Then posterior expected loss is

$$E[L_B(\tilde{\theta}, \theta)|\boldsymbol{x}] = w[\hat{\sigma}^2 + (\tilde{\theta} - \bar{x})^2] + (1-w)[(\tilde{\theta} - \hat{\theta}^B)^2 + v] \tag{1.13}$$

where $\hat{\theta}^B$ is the posterior mean and $v = E[(\theta - \hat{\theta}^B)^2|\boldsymbol{x}]$ is the posterior variance. On completing the square on $\tilde{\theta}$ in (1.13), we have

$$E[L_B(\tilde{\theta}, \theta)|\boldsymbol{x}] = w\hat{\sigma}^2 + (1-w)v + w(1-w)(\bar{x} - \hat{\theta}^B)^2 + (\tilde{\theta} - \tilde{\theta}_*)^2, \tag{1.14}$$

where

$$\tilde{\theta}_* = w\bar{x} + (1-w)\hat{\theta}^B. \tag{1.15}$$

From (1.14), it is clear that $\tilde{\theta}_*$ in (1.15) is the value of $\tilde{\theta}$ that leads to minimal posterior expected loss and is thus the Bayesian estimate of θ relative to BLF in (1.12).

Thus conditional on the data and prior information, $\tilde{\theta}_*$ in (1.15) is optimal in the sense of providing minimal posterior expected loss. If $w = 1$ in (1.15), $\tilde{\theta}_* = \bar{x}$, while if $w = 0$, $\tilde{\theta}_* = \hat{\theta}^B$.

In the multivariate case, consider the model

$$\boldsymbol{X} = \boldsymbol{\theta} + \boldsymbol{u},$$

where $\boldsymbol{\theta}^T = (\theta_1, \theta_2, ..., \theta_m)$ and $\boldsymbol{u}^T = (u_1, u_2, ..., u_m)$. The u_i^Ts are assumed to be *iid* $N(0, \sigma^2)$. Thus X_i is normally distributed with mean θ_i and variance σ^2. The problem is to estimate $\boldsymbol{\theta}$ relative to the following BLF:

$$L_B(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = w(\boldsymbol{x} - \tilde{\boldsymbol{\theta}})^T(\boldsymbol{x} - \tilde{\boldsymbol{\theta}}) + (1-w)(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta})^T(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) \tag{1.16}$$

where $\tilde{\boldsymbol{\theta}}$ is some estimate of $\boldsymbol{\theta}$ and w is a given weight, $0 \le w \le 1$. The first term on the r.h.s. of (1.16) represents goodness of fit while the second represents precision of estimation.

Thus the first problem is to find a value of $\tilde{\boldsymbol{\theta}}$ which minimizes the posterior expectation of the loss function in (1.16). Given a posterior probability density

function for $\boldsymbol{\theta}$, one finds

$$\begin{aligned} E[L_B(\tilde{\boldsymbol{\theta}},\boldsymbol{\theta})|\boldsymbol{x}] &= w(\boldsymbol{x}-\tilde{\boldsymbol{\theta}})^T(\boldsymbol{x}-\tilde{\boldsymbol{\theta}}) \\ &+ (1-w)E[\{\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B-(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^B)\}^T\{\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B-(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^B)\}|\boldsymbol{x}] \\ &= w(\boldsymbol{x}-\tilde{\boldsymbol{\theta}})^T(\boldsymbol{x}-\tilde{\boldsymbol{\theta}})+(1-w)(\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B)^T(\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B) \\ &+ (1-w)E[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^B)^T(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^B)|\boldsymbol{x}] \qquad (1.17) \end{aligned}$$

where $\hat{\boldsymbol{\theta}}^B$ is the posterior mean of $\boldsymbol{\theta}$. On completing the square on $\tilde{\boldsymbol{\theta}}$ in (1.17), we have

$$\begin{aligned} E[L_B(\tilde{\boldsymbol{\theta}},\boldsymbol{\theta})|\boldsymbol{x}] &= (\tilde{\boldsymbol{\theta}}-\tilde{\boldsymbol{\theta}}^*)^T(\tilde{\boldsymbol{\theta}}-\tilde{\boldsymbol{\theta}}^*)+w(1-w)(\boldsymbol{x}-\hat{\boldsymbol{\theta}}^B)^T(\boldsymbol{x}-\hat{\boldsymbol{\theta}}^B) \\ &+ (1-w)E[(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^B)^T(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}^B)|\boldsymbol{x}] \end{aligned}$$

where

$$\tilde{\boldsymbol{\theta}}^* = w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B.$$

From (1.17), $\tilde{\boldsymbol{\theta}}^*$ is the value of $\tilde{\boldsymbol{\theta}}$ that minimizes posterior loss.

As an example, let $X_1,...,X_m|\theta_1,...,\theta_m \sim N(\theta_i,1)$ and $\theta_i \sim N(\mu,A)$, assuming μ, A is unknown. Marginally $X_i \sim N(\mu, 1/B)$, where $B=(1+A)^{-1}$.

Under the balanced loss function defined as before, new Bayes estimators, which are not the posterior means any more, can be derived. To this end, we calculate

$$\begin{aligned} E[L_B(\tilde{\boldsymbol{\theta}},\boldsymbol{\theta})|\boldsymbol{x}] &= w||\boldsymbol{x}-\tilde{\boldsymbol{\theta}}||^2+(1-w)E[||\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B+\hat{\boldsymbol{\theta}}^B-\boldsymbol{\theta}||^2|\boldsymbol{x}] \\ &= w||\boldsymbol{x}-\tilde{\boldsymbol{\theta}}||^2 \\ &+ (1-w)\{||\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B||^2+E(||\hat{\boldsymbol{\theta}}^B-\boldsymbol{\theta}||^2|\boldsymbol{x})\} \end{aligned}$$

where w is weight, i.e., $0 \le w \le 1$, and $\hat{\boldsymbol{\theta}}^B = E(\boldsymbol{\theta}|\boldsymbol{x})$.

Completing the square on $\tilde{\boldsymbol{\theta}}$

$$\begin{aligned}
E[L_B(\tilde{\boldsymbol{\theta}},\boldsymbol{\theta})|\boldsymbol{x}] &= w(\boldsymbol{x}-\tilde{\boldsymbol{\theta}})^T(\boldsymbol{x}-\tilde{\boldsymbol{\theta}}) + (1-w)(\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B)^T(\tilde{\boldsymbol{\theta}}-\hat{\boldsymbol{\theta}}^B) \\
&+ (1-w)E[||\hat{\boldsymbol{\theta}}^B-\boldsymbol{\theta}||^2|\boldsymbol{x}] \\
&= \tilde{\boldsymbol{\theta}}^T\tilde{\boldsymbol{\theta}} - 2\tilde{\boldsymbol{\theta}}^T(w\boldsymbol{x}+\hat{\boldsymbol{\theta}}^B - w\hat{\boldsymbol{\theta}}^B) \\
&+ (w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B)^T(w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B) \\
&- (w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B)^T(w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B) + w\boldsymbol{x}^T\boldsymbol{x} + (1-w)\hat{\boldsymbol{\theta}}^{B^T}\hat{\boldsymbol{\theta}}^B \\
&+ (1-w)E[||\hat{\boldsymbol{\theta}}^B-\boldsymbol{\theta}||^2|\boldsymbol{x}] \\
&= (\tilde{\boldsymbol{\theta}}-(w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B))^T(\tilde{\boldsymbol{\theta}}-(w\boldsymbol{x}+(1-w)\hat{\boldsymbol{\theta}}^B)) \\
&+ w(1-w)(\boldsymbol{x}-\hat{\boldsymbol{\theta}}^B)^T(\boldsymbol{x}-\hat{\boldsymbol{\theta}}^B) + (1-w)E[||\hat{\boldsymbol{\theta}}^B-\boldsymbol{\theta}||^2|\boldsymbol{x}].
\end{aligned}$$

So, our new Bayes and empirical Bayes estimators under balanced loss function is as follows:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{NB} &= w\boldsymbol{X} + (1-w)\hat{\boldsymbol{\theta}}^B = w\boldsymbol{X} + (1-w)[(1-B)\boldsymbol{X} + B\mu\mathbf{1}] \\
&= [1-(1-w)B]\boldsymbol{X} + (1-w)B\mu\mathbf{1}
\end{aligned}$$

$$\hat{\boldsymbol{\theta}}^{NEB} = [1-(1-w)\hat{B}]\boldsymbol{X} + (1-w)\hat{B}\bar{X}\mathbf{1}$$

where $\hat{B} = (m-3)/\sum(X_i-\bar{X})^2,\ m \geq 4$.

Note that in case of $w = 0$, it is an empirical Bayes estimator under quadratic loss function.

1.2 The Subject of This Dissertation

In Chapter 2, we consider the asymptotic expansion of the MSE of constrained James-Stein estimators. This expansion is valid up to $O(m^{-1})$. We also provide an estimator of the MSE which is asymptotically valid up to $O(m^{-1})$, m denoting the number of strata. A simulation study is undertaken to evaluate the performance of these estimators.

Chapter 3 develops constrained Bayes and empirical Bayes estimators under balanced loss functions. In particular, such estimators are derived under the one-parameter exponential family of distributions. In the normal-normal example, asymptotic expansions of MSE's of the Bayes and empirical Bayes estimators are provided which are asymptotically valid up to $O(m^{-1})$. In addition, similar asymptotic expansions of MSE's of constrained Bayes and empirical Bayes estimators are also provided. Estimators of these MSE's asymptotically valid up to $O(m^{-1})$ are also provided.

Chapter 4 develops constrained Bayes and constrained empirical Bayes estimators for the random effects balanced normal ANOVA model when both variance components are unknown. The asymptotic MSE's valid up to $O(m^{-1})$ are derived as in the previous chapters.

Finally, in Chapter 5, we summarize the result of this dissertation and propose several topics for future research.

CHAPTER 2
ASYMPTOTIC MEAN SQUARED ERRORS AND ESTIMATION

2.1 Introduction

James-Stein estimators (James and Stein, 1961) have long been popular among statisticians. The theoretical interest in these estimators stems from their minimaxity and other related properties. On the other hand, practitioners have found these estimators quite appealing in the context of simultaneous estimation of parameters when there is a clear need for borrowing strength. While the original James-Stein estimators shrink the multivariate sample mean towards some prior mean, Lindley's (1962) modification of the same shrinks the sample mean towards some grand average. All these estimators have interesting empirical Bayes (EB) interpretation (see Efron and Morris, 1973).

However, as discussed in the previous chapter, the histogram of the posterior means of co-ordinate specific parameters is underdispersed as an estimate of the parameter histogram. The EB estimators, usually derived from the posterior means by plugging in estimators of the hyperparameters, share the same feature. It is thus clear that with the twin objective of simultaneous estimation of parameters under the quadratic loss, and achieving closeness of the histogram of the posterior means with the posterior estimate of the parameter histogram, the usual Bayes or EB estimators are clearly inappropriate. Indeed, any single set of values cannot simultaneously optimize the two goals. However, in many policy settings reporting a "single" set of estimates with "good" performance for these two goals is a clear necessity.

Louis (1984) addressed this problem by matching the first two empirical moments of the Bayes estimates of the normal means with the corresponding

moments derived from the posterior histogram. The Bayes estimators which meet these constraints are referred to as constrained Bayes (CB) estimators. Louis proposed also constrained empirical Bayes (CEB) estimators in the original James-Stein framework. Ghosh (1992) developed CB estimators in a more general framework when the distribution was not necessarily normal.

As a natural next step, one needs to find measures of precision associated with the CEB estimators. Reporting a set of estimates without any associated measures of uncertainty is against standard statistical practice. Indeed, very often measures of precision along with the estimators are demanded by users of the data. For example, in finding asymptotic confidence sets for the parameter vector of interest centered at the CEB estimators, one needs at least, some asymptotic approximation of its mean squared error (MSE). Throughout, we will use the term MSE as equivalent to the Bayes risk. Unlike the regular James-Stein estimators, it seems impossible to find exact MSE's of the CEB estimators. We provide instead the asymptotic MSE's of these estimators which are correct up to order $O(m^{-1})$. We find also bias-corrected estimators of these MSE's which are also asymptotically valid up to $O(m^{-1})$. Our results are thus similar in spirit to those of Prasad and Rao (1990), Lahiri and Rao (1995) and Datta and Lahiri (2000).

We may point out that when the Bayesian model is true, the Bayes estimators have the smallest Bayes risks. The CB estimators cannot claim any risk improvement over the Bayes estimators. The same phenomenon is reflected in the comparison of EB and CEB estimators. The CEB estimators are not designed to improve on the EB estimators by producing smaller MSE's. They are constructed to meet the twin objectives as mentioned earlier in this section more satisfactorily than the EB estimators.

In Section 2 of this chapter, we provide the asymptotic expansion of the MSE which is correct up to $O(m^{-1})$. Section 3 contains estimators of the MSE which is

also asymptotically valid up to $O(m^{-1})$. Section 4 contains some simulation results demonstrating the accuracy of all the approximations.

2.2 The Asymptotic MSE

Consider the usual normal-normal model where $X_i|\theta_i$ are independent $\mathrm{N}(\theta_i, 1)$, while θ_i are iid $\mathrm{N}(\mu, A)$. Let $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_m)^T$. For an estimate $\boldsymbol{d} = (d_1, \cdots, d_m)^T$ of $\boldsymbol{\theta}$, consider the loss $\mathrm{L}(\boldsymbol{\theta}, \boldsymbol{d}) = m^{-1}||\boldsymbol{\theta} - \boldsymbol{d}||^2$, where $||\cdot||$ denotes the Euclidean norm. Then the Bayes estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}^B = (\hat{\theta}_1^B(\boldsymbol{X}), ..., \hat{\theta}_m^B(\boldsymbol{X}))^T,$$

where $\hat{\theta}_i^B(\boldsymbol{X}) = (1-B)X_i + B\mu$, $i = 1, \cdots, m$, $B = (1+A)^{-1}$.

2.2.1 Constrained James-Stein Estimators

Following Louis (1984) and Ghosh (1992), the CB estimator of $\boldsymbol{\theta}$ is given by

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{CB} &= a_B[(1-B)\boldsymbol{X} + B\mu\mathbf{1}_m] + (1-a_B)[(1-B)\bar{X} + B\mu]\mathbf{1}_m \\ &= (1-B)[a_B\boldsymbol{X} + (1-a_B)\bar{X}\mathbf{1}_m] + B\mu\mathbf{1}_m, \end{aligned} \tag{2.1}$$

where $\mathbf{1}_m$ is an m-component column vector with each element equal to 1, $\bar{X} = m^{-1}\sum_{i=1}^m X_i$, $a_B^2 = 1 + \frac{1}{(1-B)S}$, and $S = \sum_{i=1}^m (X_i - \bar{X})^2/(m-1)$.

In an EB scenario, typically both μ and A are unknown, and are estimated from the marginal distribution of $\boldsymbol{X} = (X_1, \cdots, X_m)^T$. Marginally, $\boldsymbol{X} \sim N(\mu\mathbf{1}_m, B^{-1}\boldsymbol{I}_m)$. Hence, marginally, $S \sim B^{-1}\chi^2_{m-1}/(m-1)$. We estimate μ by $\bar{X}$, and B by $\hat{B} = \min(\frac{m-d}{m-1}, \frac{1}{S})$, where $d > 1$. For our asymptotic MSE expansion, any $d > 1$ should do. Morris (1981) takes $d = 3$. The constrained EB estimator of $\boldsymbol{\theta}$ is then given by

$$\hat{\boldsymbol{\theta}}^{CEB} = (1-\hat{B})[a_{EB}\boldsymbol{X} + (1-a_{EB})\bar{X}\mathbf{1}_m] + \hat{B}\bar{X}\mathbf{1}_m, \tag{2.2}$$

where a_{EB} replaces B by $\hat{B}$ in a_B.

Under the assumed loss, the MSE of $\hat{\boldsymbol{\theta}}^{CEB}$ is given by

$$\begin{aligned}
\mathrm{MSE}(\hat{\boldsymbol{\theta}}^{CEB}) &= m^{-1}E\|\hat{\boldsymbol{\theta}}^{CEB}-\boldsymbol{\theta}\|^2 \\
&= m^{-1}E\|\hat{\boldsymbol{\theta}}^{CEB}-\hat{\boldsymbol{\theta}}^{B}+\hat{\boldsymbol{\theta}}^{B}-\boldsymbol{\theta}\|^2 \\
&= m^{-1}[E\|\hat{\boldsymbol{\theta}}^{CEB}-\hat{\boldsymbol{\theta}}^{B}\|^2+E\|\hat{\boldsymbol{\theta}}^{B}-\boldsymbol{\theta}\|^2]. \qquad (2.3)
\end{aligned}$$

The product term disappears since $E(\hat{\boldsymbol{\theta}}^{B}-\boldsymbol{\theta}|\boldsymbol{X})=\mathbf{0}$.

It is well-known that

$$m^{-1}E\|\hat{\boldsymbol{\theta}}^{B}-\boldsymbol{\theta}\|^2 = m^{-1}\sum_{i=1}^{m}E(\hat{\theta}_i^B-\theta_i)^2 = 1-B. \qquad (2.4)$$

The following theorem provides an asymptotic expansion of $m^{-1}E\|\hat{\boldsymbol{\theta}}^{CEB}-\hat{\boldsymbol{\theta}}^{B}\|^2$ correct up to $O(m^{-1})$.

<u>Theorem 2.1</u> Under the given model and the loss,

$$\begin{aligned}
m^{-1}E\|\hat{\boldsymbol{\theta}}^{CEB}-\hat{\boldsymbol{\theta}}^{B}\|^2 &= \frac{B}{m}+\frac{(m-1)}{m}\frac{(1-B)}{B}[1-(1-B)^{1/2}]^2+\frac{1}{2m}B(1-B)^{-1/2} \\
&+ o(m^{-1}). \qquad (2.5)
\end{aligned}$$

<u>Remark 2.1</u> Combining (2.3)-(2.5), one gets

$$\begin{aligned}
\mathrm{MSE}(\hat{\boldsymbol{\theta}}^{CEB}) &= 1-B+B^{-1}(1-B)[1-(1-B)^{1/2}]^2 \\
&+m^{-1}[B-B^{-1}(1-B)\{1-(1-B)^{1/2}\}^2+(1/2)B(1-B)^{-1/2}] \\
&+o(m^{-1}). \qquad (2.6)
\end{aligned}$$

<u>Proof of Theorem 2.1</u> First, by (2.1) and (2.2) we get

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{CEB}-\hat{\boldsymbol{\theta}}^{B} &= (1-\hat{B})[a_{EB}\boldsymbol{X}+(1-a_{EB})\bar{X}\mathbf{1}_m]+\hat{B}\bar{X}\mathbf{1}_m-(1-B)\boldsymbol{X}-B\mu\mathbf{1}_m \\
&= a_{EB}(1-\hat{B})(\boldsymbol{X}-\bar{X}\mathbf{1}_m)+\bar{X}\mathbf{1}_m-(1-B)(\boldsymbol{X}-\bar{X}\mathbf{1}_m) \\
&- (1-B)\bar{X}\mathbf{1}_m-B\mu\mathbf{1}_m \\
&= B(\bar{X}-\mu)\mathbf{1}_m+[a_{EB}(1-\hat{B})-(1-B)](\boldsymbol{X}-\bar{X}\mathbf{1}_m). \qquad (2.7)
\end{aligned}$$

From the independence of $\bar{X}$ and $\boldsymbol{X} - \bar{X}\mathbf{1}_m$, and the fact that $\bar{X} \sim N(\mu, (mB)^{-1})$, we get from (2.7)

$$m^{-1}E||\hat{\boldsymbol{\theta}}^{CEB} - \hat{\boldsymbol{\theta}}^{B}||^2 = \frac{B}{m} + \frac{m-1}{m}E[\{a_{EB}(1-\hat{B}) - (1-B)\}^2 S]. \tag{2.8}$$

Let $g(S) = [a_{EB}(1-\hat{B}) - (1-B)]^2 S$. We write $g(S) = g_1(S) + g_2(S)$, where $g_1(S) = g(S)I_{[S>1+\epsilon_m]}$ and $g_2(S) = g(S)I_{[S\leq 1+\epsilon_m]}$, where $\epsilon_m = O(m^{-\alpha})$, and $\alpha \in (0,1)$ will be chosen later. Next we use the inequality,

$$\begin{aligned} g_2(S) &\leq 2[a_{EB}^2(1-\hat{B})^2 + (1-B)^2]SI_{[S\leq 1+\epsilon_m]} \\ &= 2[(1-\hat{B})^2 S + 1 - \hat{B} + (1-B)^2 S]I_{[S\leq 1+\epsilon_m]} \\ &\leq 2[S + 1 + S]I_{[S\leq 1+\epsilon_m]} \\ &\leq 2[2(1+\epsilon_m) + 1]I_{[S\leq 1+\epsilon_m]}. \end{aligned}$$

Hence,

$$E[g_2(S)] \leq 2[2(1+\epsilon_m) + 1]P(S \leq 1 + \epsilon_m). \tag{2.9}$$

Since $S \sim B^{-1}\chi^2_{m-1}/(m-1)$ so that $E(S) = B^{-1}$, we get

$$\begin{aligned} P(S \leq 1 + \epsilon_m) &= P(S - B^{-1} \leq 1 - B^{-1} + \epsilon_m) \\ &\leq P(|S - B^{-1}| \geq B^{-1} - 1 - \epsilon_m), \end{aligned} \tag{2.10}$$

where m is taken sufficiently large so that $B^{-1} - 1 - \epsilon_m > 0$. Now by Markov's inequality,

$$P(|S - B^{-1}| \geq B^{-1} - 1 - \epsilon_m) \leq \frac{E|S - B^{-1}|^r}{(B^{-1} - 1 - \epsilon_m)^r} = O(m^{-r/2}) \tag{2.11}$$

for $r > 0$. Choose $r > 2$ so that the right hand side of (2.11) is $O(m^{-1})$. Combining (2.9)-(2.11),

$$E[g_2(S)] = o(m^{-1}). \tag{2.12}$$

Next observe that when $S > 1+\epsilon_m$, $S^{-1} \le (1+\epsilon_m)^{-1} = [1+O(m^{-\alpha})]^{-1}$, while $(m-d)/(m-1) = [1+(d-1)/(m-d)]^{-1} = [1+O(m^{-1})]^{-1}$.

Since for large m, $1+O(m^{-\alpha}) > 1+O(m^{-1})$, $S^{-1} < (m-d)/(m-1)$ for large m. Hence, for large m, $\hat{B} = S^{-1}$. Thus,

$$\begin{aligned} g_1(S) &= [(\frac{S}{S-1})^{1/2}(\frac{S-1}{S}) - (1-B)]^2 S I_{[S>1+\epsilon_m]} \\ &= [S-1+(1-B)^2 S - 2(1-B)S^{1/2}(S-1)^{1/2}] I_{[S>1+\epsilon_m]} \\ &= h(S) I_{[S>1+\epsilon_m]} \quad \text{say,} \end{aligned}$$

where

$$h(x) = x-1+(1-B)^2x - 2(1-B)x^{1/2}(x-1)^{1/2}, \quad x>1. \tag{2.13}$$

By Taylor expansion again,

$$\begin{aligned} h(S) &= h(ES) + (S-ES)h'(ES) + \frac{1}{2}(S-ES)^2 h''(ES) \\ &+ \frac{1}{2}\int_{ES}^{S}(S-x)^2 h'''(x)dx, \end{aligned} \tag{2.14}$$

where by (2.13), for $x > 1$,

$$h'(x) = 1+(1-B)^2 - (1-B)\{x^{-1/2}(x-1)^{1/2} + x^{1/2}(x-1)^{-1/2}\};$$

$$\begin{aligned} h''(x) &= \frac{1-B}{2}\{x^{-3/2}(x-1)^{1/2} - 2x^{-1/2}(x-1)^{-1/2} + x^{1/2}(x-1)^{-3/2}\} \\ &= \frac{1-B}{2}x^{-3/2}(x-1)^{-3/2}; \end{aligned}$$

$$h'''(x) = -\frac{3(1-B)}{4}x^{-5/2}(x-1)^{-5/2}(2x-1).$$

On simplification,

$$\begin{aligned} h(ES) = h(B^{-1}) &= B^{-1}-1+(1-B)^2B^{-1} - 2(1-B)B^{-1/2}(B^{-1}-1)^{-1/2} \\ &= B^{-1}(1-B)[1-(1-B)^{1/2}]^2; \end{aligned} \tag{2.15}$$

$$h''(ES) = h''(B^{-1}) = \frac{1}{2}B^3(1-B)^{-1/2}. \tag{2.16}$$

Note also that $E[h(ES)I_{[S\leq 1+\epsilon_m]}] = o(m^{-1})$ and by the Schwarz inequality

$$\begin{aligned}
E|(S-ES)h'(ES)I_{[S\leq 1+\epsilon_m]}| &= |h'(ES)|E|(S-ES)I_{[S\leq 1+\epsilon_m]}| \\
&\leq |h'(B^{-1})|E^{1/2}(S-ES)^2P^{1/2}(S\leq 1+\epsilon_m) \\
&= |h'(B^{-1})|O(m^{-1/2})O(m^{-r/4}) \\
&= o(m^{-1}), \qquad (2.17)
\end{aligned}$$

by choosing $r > 2$. Also,

$$\begin{aligned}
E[(S-ES)^2h''(ES)I_{[S\leq 1+\epsilon_m]}] &= h''(B^{-1})E[(S-ES)^2I_{[S\leq 1+\epsilon_m]}] \\
&\leq h''(B^{-1})E^{1/2}(S-ES)^4P^{1/2}(S\leq 1+\epsilon_m) \\
&= O(m^{-1-r/4}) = o(m^{-1}) \quad \text{for} \quad r>0. \qquad (2.18)
\end{aligned}$$

Noting $h(ES) = h(ES)I_{[S>1+\epsilon_m]} + h(ES)I_{[S\leq 1+\epsilon_m]} = h(ES)I_{[S>1+\epsilon_m]} + o(m^{-1})$, and similarly $h'(ES) = h'(ES)I_{[S>1+\epsilon_m]} + o(m^{-1})$, $h''(ES) = h(ES)I_{[S>1+\epsilon_m]} + o(m^{-1})$, it follows from (2.14)-(2.18) that

$$\begin{aligned}
E[g_1(S)] &= E[h(S)I_{[S>1+\epsilon_m]}] \\
&= h(ES) + E(S-ES)h'(ES) + \frac{1}{2}E(S-ES)^2h''(ES) \\
&\quad + \frac{1}{2}E[\int_{ES}^{S}(S-x)^2h'''(x)dxI_{[S>1+\epsilon_m]}] + o(m^{-1}) \\
&= B^{-1}(1-B)[1-(1-B)^{1/2}]^2 + \frac{1}{2}\frac{2}{(m-1)B^2}\frac{1}{2}B^3(1-B)^{-1/2} \\
&\quad + \frac{1}{2}E[\int_{ES}^{S}(S-x)^2h'''(x)dxI_{[S>1+\epsilon_m]}] + o(m^{-1}) \\
&= B^{-1}(1-B)[1-(1-B)^{1/2}]^2 + \frac{1}{2m}B(1-B)^{-1/2} \\
&\quad + \frac{1}{2}E[\int_{ES}^{S}(S-x)^2h'''(x)dxI_{[S>1+\epsilon_m]}] + o(m^{-1}).
\end{aligned}$$

Finally, since $S > 1+\epsilon_m$ and $E(S) = B^{-1} > 1+\epsilon_m$ for large m,

$$
\begin{aligned}
E[|\int_{ES}^{S}(S-x)^2 h'''(x)dx|I_{[S>1+\epsilon_m]}] &\le \frac{3(1-B)}{4}\epsilon_m^{-5/2}2E[|\int_{ES}^{S}(S-x)^2dx|I_{[S>1+\epsilon_m]}] \\
&\le \frac{3(1-B)}{2}\epsilon_m^{-5/2}E|S-ES|^3 \\
&= O(m^{5\alpha/2-3/2}) = o(m^{-1}) \qquad (2.19)
\end{aligned}
$$

if $\frac{5\alpha-3}{2} < -1$, i.e., $\alpha < 1/5$. Combining (2.8), (2.12) and (2.19), one gets (2.5).

2.2.2 Positive-Part James-Stein Estimators

We first observe that

$$
\hat{\boldsymbol{\theta}}^{EB} = (1-\hat{B})\boldsymbol{X} + \hat{B}\bar{X}\mathbf{1}_m. \qquad (2.20)
$$

Here we may recall that $\hat{B} = \min(\frac{m-d}{m-1}, \frac{1}{S})$ and $S = \sum_{i=1}^{m}(X_i - \bar{X})^2/(m-1)$. Then, under the assumed loss, the MSE of $\hat{\boldsymbol{\theta}}^{EB}$ is given by

$$
\begin{aligned}
\mathrm{MSE}(\hat{\boldsymbol{\theta}}^{EB}) &= m^{-1}E||\hat{\boldsymbol{\theta}}^{EB} - \boldsymbol{\theta}||^2 \\
&= m^{-1}E||\hat{\boldsymbol{\theta}}^{EB} - \hat{\boldsymbol{\theta}}^{B} + \hat{\boldsymbol{\theta}}^{B} - \boldsymbol{\theta}||^2 \\
&= m^{-1}[E||\hat{\boldsymbol{\theta}}^{EB} - \hat{\boldsymbol{\theta}}^{B}||^2 + E||\hat{\boldsymbol{\theta}}^{B} - \boldsymbol{\theta}||^2] \\
&= 1 - B + m^{-1}[E||\hat{\boldsymbol{\theta}}^{EB} - \hat{\boldsymbol{\theta}}^{B}||^2]
\end{aligned}
$$

The following theorem provides an asymptotic expansion of MSE of $\hat{\boldsymbol{\theta}}^{EB}$ correct up to $O(m^{-1})$.

Theorem 2.2 Under the given model and the loss,

$$
m^{-1}E||\hat{\boldsymbol{\theta}}^{EB} - \boldsymbol{\theta}^{B}||^2 = 1 - B + \frac{3B}{m} + o(m^{-1}).
$$

Proof of Theorem 2.2 First, by (2.1) and (2.20) we get

$$
\begin{aligned}
\hat{\boldsymbol{\theta}}^{EB} - \hat{\boldsymbol{\theta}}^{B} &= (1-\hat{B})\boldsymbol{X} + \hat{B}\bar{X}\mathbf{1}_m - (1-B)\boldsymbol{X} - B\mu\mathbf{1}_m \\
&= (B-\hat{B})(\boldsymbol{X} - \bar{X}\mathbf{1}_m) + B(\bar{X}-\mu)\mathbf{1}_m.
\end{aligned}
$$

After some simplification,

$$m^{-1}E||\hat{\boldsymbol{\theta}}^{EB} - \boldsymbol{\theta}||^2 = 1 - B + Bm^{-1} + \frac{m-1}{m}E[(\hat{B} - B)^2 S]. \tag{2.21}$$

Writing once again $\hat{\hat{B}} = S^{-1}$, and by $E[S] = B^{-1}$,

$$\begin{aligned} E[(\hat{\hat{B}} - B)^2 S] &= E[S^{-1} - 2B + B^2 S] \\ &= \frac{(m-1)B}{m-3} + B \\ &= 2B(m-3)^{-1} \text{ for } m \geq 4, \end{aligned} \tag{2.22}$$

and it follows after some simplifications and calculations with same technique as before,

$$E[(\hat{\hat{B}} - \hat{B})^2 S] = o(m^{-1}). \tag{2.23}$$

Combining (2.21)-(2.23), it follows now that MSE of $\hat{\boldsymbol{\theta}}^{EB}$ correct up to $O(m^{-1})$ is given by $1 - B + \frac{3B}{m}$.

The next section will be devoted to asymptotic estimation of the MSE.

2.3 Estimation of the MSE

This section is devoted to estimation of the MSE derived in Section 2. We find estimators which are asymptotically correct up to $O(m^{-1})$.

2.3.1 Constrained James-Stein Estimators

First, by (2.6) we express $m^{-1}E||\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{\theta}||^2$ as

$$m^{-1}E||\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{\theta}||^2 = w_1(B) + m^{-1}w_2(B) + o(m^{-1}), \tag{2.24}$$

where $w_1(B) = 1 - B + A(B)$, $A(B) = B^{-1}(1-B)[1-(1-B)^{1/2}]^2 = 2B^{-1} - 3 + B - 2B^{-1}(1-B)^{3/2}$, and $w_2(B) = B - A(B) + C(B)$, $C(B) = (1/2)B(1-B)^{-1/2}$.

The following theorem provides estimators of the MSE of constrained James-Stein estimators which is asymptotically correct up to $O(m^{-1})$.

Theorem 2.3 The $O(m^{-1})$ bias-corrected estimator of the MSE of $\hat{\theta}^{CEB}$ is given by

$$w_1(\hat{B}) + m^{-1}[w_2(\hat{B}) + (3/2)C(\hat{B})],$$

where $w_1(B)$, $w_2(B)$ and $C(B)$ are defined after (2.24).

Proof of Theorem 2.3 We begin with

$$E[(1-\hat{B})I_{[S\le 1+\epsilon_m]}] \le P(S \le 1+\epsilon_m) = O(m^{-r}) = o(m^{-1}) \quad \text{for} \quad r > 2;$$

$$E[(1-\hat{B})I_{[S>1+\epsilon_m]}] = E[(1-S^{-1})I_{[S>1+\epsilon_m]}] = E[g_0(S)I_{[S>1+\epsilon_m]}],$$

where $g_0(x) = 1 - x^{-1}$, $g_0'(x) = x^{-2}$, $g_0''(x) = -2x^{-3}$, and $g_0'''(x) = 6x^{-4}$. By Taylor expansion,

$$\begin{aligned} g_0(S) &= g_0(ES) + (S-ES)g_0'(ES) + (1/2)(S-ES)^2 g_0''(ES) \\ &\quad + (1/2)\int_{ES}^{S}(S-x)^2 g_0'''(x)dx \end{aligned}$$

where $g_0(ES) = 1 - B$, $g_0''(ES) = -2B^3$

Also note that for $x > 1$, $|g_0'''(x)| \le \frac{6}{x^4} \le 6$. Hence, by the same argument as in the previous section,

$$\begin{aligned} E(1-\hat{B}) &= 1 - B + (1/2)2[(m-1)B^2]^{-1}(-2B^3) + o(m^{-1}) \\ &= 1 - B - 2Bm^{-1} + o(m^{-1}) \end{aligned} \tag{2.25}$$

Next note that

$$\begin{aligned} A(\hat{B})I_{[S\le 1+\epsilon_m]} &\le \hat{B}^{-1}I_{[S\le 1+\epsilon_m]} \\ &= \max((m-1)/(m-d), S)I_{[S\le 1+\epsilon_m]} \\ &\le (1+\epsilon_m)I_{[S\le 1+\epsilon_m]}, \end{aligned}$$

since $(m-1)/(m-d) = 1 + O(m^{-1}) \le 1 + \epsilon_m$. Hence, for large m,

$$E[A(\hat{B})I_{[S\le 1+\epsilon_m]}] \le (1+\epsilon_m)P(S \le 1+\epsilon_m) \le O(m^{-r/2}) = o(m^{-1})$$

for $r > 2$ and since $\hat{B} = S^{-1}$ for $S^{-1} < \frac{m-1}{m-d}$, for such m,

$$\begin{aligned} A(\hat{B})I_{[S>1+\epsilon_m]} &= [2S - 3 + \frac{1}{S} - \frac{2(S-1)^{3/2}}{S^{1/2}}]I_{[S>1+\epsilon_m]} \\ &= u(S)I_{[S>1+\epsilon_m]}, \end{aligned}$$

where

$$u(x) = 2x - 3 + \frac{1}{x} - 2(x-1)^{3/2}x^{-1/2}.$$

By Taylor expansion again,

$$\begin{aligned} u(S) &= u(ES) + (S-ES)u'(ES) + (1/2)(S-ES)^2u''(ES) \\ &\quad + (1/2)\int_{ES}^{S}(S-x)^2u'''(x)dx, \end{aligned}$$

where

$$u'(x) = 2 - \frac{1}{x^2} - 3(1-\frac{1}{x})^{1/2} + (1-\frac{1}{x})^{3/2};$$

$$u''(x) = \frac{2}{x^3} - \frac{3}{2}(1-\frac{1}{x})^{-1/2}(\frac{1}{x^2}) + \frac{3}{2}(1-\frac{1}{x})^{1/2}(\frac{1}{x^2});$$

$$\begin{aligned} u'''(x) &= -\frac{6}{x^4} + \frac{3}{4}(1-\frac{1}{x})^{-3/2}(\frac{1}{x^4}) + \frac{3}{4}(1-\frac{1}{x})^{-1/2}(\frac{1}{x^4}) \\ &\quad - 3(1-\frac{1}{x})^{1/2}(\frac{1}{x^3}) + 3(1-\frac{1}{x})^{-1/2}(\frac{1}{x^3}) \\ &= -\frac{6}{x^4} + \frac{3}{4}(1-\frac{1}{x})^{-3/2}(\frac{1}{x^4}) + \frac{3}{4}(1-\frac{1}{x})^{-1/2}(\frac{1}{x^4}) \\ &\quad + 3(\frac{1}{x^{7/2}})(x-1)^{-1/2} \\ &= -\frac{6}{x^4} + \frac{3}{4}(\frac{1}{x^{5/2}})(\frac{1}{(x-1)^{3/2}}) + \frac{15}{4}(\frac{1}{x^{7/2}})(\frac{1}{(x-1)^{1/2}}). \end{aligned}$$

On simplification,

$$u(ES) = u(B^{-1}) = 2B^{-1} - 3 + B - 2B^{-1}(1-B)^{3/2};$$

$$u''(ES) = u''(B^{-1}) = 2B^3 - \frac{3}{2}B^3(1-B)^{-1/2};$$

and $|u'''(x)| \leq \frac{6}{(1+\epsilon_m)^4} + \frac{3\epsilon_m^{-3/2}}{4(1+\epsilon_m)^{5/2}} + \frac{15\epsilon_m^{-1/2}}{4(1+\epsilon_m)^{7/2}} = a_m$ (say). Hence,

$$\begin{aligned} E[|\int_{ES}^{S}(S-x)^2u'''(x)dx|I_{[S>1+\epsilon_m]}] &\leq a_m E|S-ES|^3 \\ &= O(m^{3\alpha/2-3/2}) = o(m^{-1}) \end{aligned}$$

for $\alpha < 1/3$. Thus,

$$E[A(\hat{B})] = A(B) + 2Bm^{-1} - (3/2)m^{-1}B(1-B)^{-1/2} + o(m^{-1}). \tag{2.26}$$

Next

$$E[\hat{B}(1-\hat{B})^{-1/2}I_{[S\leq 1+\epsilon_m]}] \leq E[(1-\hat{B})^{-1/2}I_{[S\leq 1+\epsilon_m]}]. \tag{2.27}$$

Since $1-\hat{B} \geq (d-1)/(m-1)$, right hand side of (2.27) is less than or equal to $\{(m-1)^{1/2}/(d-1)^{1/2}\}P(S \leq 1+\epsilon_m) = O(m^{1/2-r/2}) = o(m^{-1})$ by choosing $r > 3$. Again,

$$\begin{aligned} \hat{B}(1-\hat{B})^{-1/2}I_{[S>1+\epsilon_m]} &= \frac{1}{S}(1-\frac{1}{S})^{-1/2}I_{[S>1+\epsilon_m]} \\ &= S^{-1/2}(S-1)^{-1/2}I_{[S>1+\epsilon_m]} \\ &= q(S)I_{[S>1+\epsilon_m]}, \end{aligned}$$

where $q(x) = x^{-1/2}(x-1)^{-1/2}$. By Taylor expansion,

$$q(S) = q(ES) + \int_{ES}^{S} q'(x)dx,$$

where $q(ES) = q(B^{-1}) = B(1-B)^{-1/2}$ and $q'(x) = -\frac{2x-1}{2x^{3/2}(x-1)^{3/2}}$ for $x > 1+\epsilon_m$. Thus,

$$|q'(x)| \leq \frac{1}{x^{1/2}(x-1)^{3/2}} \leq \frac{1}{\epsilon_m^{3/2}}$$

for $x > 1+\epsilon_m$. Finally,

$$E[q(S)I_{[S>1+\epsilon_m]}] = B(1-B)^{-1/2} + E[\{\int_{ES}^{S} q'(x)dx\}I_{[S>1+\epsilon_m]}] + o(m^{-1})$$

and

$$E[\{\int_{ES}^{S} q'(x)dx\}I_{[S>1+\epsilon_m]}] \leq \epsilon_m^{-3/2}E|S-ES| = O(m^{3\alpha/2-1/2}) = o(1)$$

for $\alpha < \frac{1}{3}$. Hence,

$$E[C(\hat{B})] = C(B) + o(1). \tag{2.28}$$

Combining (2.24),(2.25),(2.26) and (2.28), it follows that

$$E[w_1(\hat{B}) + m^{-1}w_2(\hat{B})] = w_1(B) + m^{-1}w_2(B) - (3/2)m^{-1}C(B) + o(m^{-1}).$$

2.3.2 Positive-Part James-Stein Estimators

Now we consider the estimators of the MSE of positive-part James-Stein estimators derived in section 2. As same argument with constrained James-Stein estimators, the following theorem provides estimators of the MSE of parsitive-part James-Stein estimators which is asymptotically correct up to $O(m^{-1})$.

Theorem 2.4 The $O(m^{-1})$ bias-corrected estimator of the MSE of $\hat{\theta}^{EB}$ is given by

$$1 - \hat{B} + 5m^{-1}\hat{B}.$$

Proof of Theorem 2.4 By (2.25), we know

$$E[\hat{B}] = B + \frac{2B}{m} + o(m^{-1}), \tag{2.29}$$

so we can get

$$E[1-\hat{B}+3m^{-1}\hat{B}] \;=\; 1-B+\frac{3B}{m}-\frac{2B}{m}+o(m^{-1}) \tag{2.30}$$

Combining (2.29) and (2.30), it follows that

$$E[1-\hat{B}+5m^{-1}\hat{B}]=1-B+\frac{3B}{m}.$$

2.4 Numerical Calculations

In this section, we report the results of a simulation study to demonstrate the accuracy of the MSE approximation as described in the previous section. For the sake of comparison, we consider also the approximate estimator of the MSE of the the positive-part James-Stein estimator (the usual EB of $\boldsymbol{\theta}$).

2.4.1 Method

We now discuss the simulation. For illustration, we consider a simple normal-normal model with $\mu=0$. We investigate the performance of the simulated MSE corresponding to (2.3) as well as the asymptotically estimated MSE of the CEB estimators for several m. The simulated MSE's of the EB estimators are also calculated. Different values of $A=1,\ 2,\ 3$ are considered.

Details of our simulation study are described below.

(a) First we generate θ_i $(i=1,\cdots,m)$ from the N(0, A) distribution with fixed A value.

(b) For given θ_i $(i=1,\cdots,m)$, we generate the data $x_i,\ i=1,\cdots,m$ from the N(θ_i, 1) distribution. We repeat steps (a) and (b) $R=10,000$ times. Then we calculate EB and CEB estimates for each simulated data set.

(c) Finally we compute the simulated MSE's

$$(mR)^{-1}\sum_{i=1}^{m}\sum_{r=1}^{R}(\hat{\theta}_{ir}^{CEB}-\theta_{ir})^2 \text{ and } (mR)^{-1}\sum_{i=1}^{m}\sum_{r=1}^{R}(\hat{\theta}_{ir}^{EB}-\theta_{ir})^2$$

for different values of m after $R = 10,000$ repetitions of the experiment. In addition, we calculate asymptotically estimated MSE of the CEB estimates for the same m values.

2.4.2 Result

Table 2.1 reports the values of the simulated MSE of EB and CEB estimates as well as the asymptotically estimated MSE of the CEB estimates for $m = 10, 30, 50, 100, 300$ and for selected values of A. Since results for different values of d are similar, only the case $d = 3$ is reported. Not surprisingly, Table 2.1 shows that the simulated MSE and asymptotic MSE for the CEB estimates are fairly close even for $m = 50$. Also, the simulated MSE's of the EB estimates are relatively smaller than the simulated MSE's of the CEB estimates. The intuitive reason behind this is that the MSE of an EB estimator is asymptotically close to the MSE (or equivalently the Bayes risk) of the regular Bayes estimator, while the MSE of a constrained EB is asymptotically close to the MSE of a constrained Bayes estimator, and the latter has clearly larger Bayes risk than that of the regular Bayes estimator. We note also the first order optimality of the EB estimator noting that its Bayes risk tends to $1 - (1 + A)^{-1}$ as $m \to \infty$.

Table 2–1: Simulated MSE's of EB and CEB estimates as well as asymptotic MSE of EB and CEB estimates for selected values of A and m

A	m	MSE			
		Simulated (EB)	Asymptotic (EB)	Simulated (CEB)	Asymptotic (CEB)
1	10	0.6040	0.6500	0.6643	0.6627
	30	0.5479	0.5500	0.6123	0.6114
	50	0.5297	0.5300	0.6006	0.6011
	100	0.5147	0.5150	0.5930	0.5935
	300	0.5046	0.5050	0.5879	0.5884
2	10	0.7582	0.7667	0.7839	0.7810
	30	0.6988	0.7000	0.7479	0.7497
	50	0.6858	0.6867	0.7415	0.7434
	100	0.6760	0.6767	0.7378	0.7387
	300	0.6694	0.6700	0.7349	0.7356
3	10	0.8263	0.8250	0.8411	0.8379
	30	0.7732	0.7750	0.8127	0.8152
	50	0.7637	0.7650	0.8084	0.8107
	100	0.7567	0.7575	0.8062	0.8072
	300	0.7518	0.7525	0.8043	0.8050

CHAPTER 3
CONSTRAINED BAYES AND EMPIRICAL BAYES ESTIMATION WITH BALANCED LOSS FUNCTION

3.1 Introduction

Constrained Bayes and empirical Bayes estimators are developed in this chapter under balanced loss. As mentioned in the introduction, such losses are formulated to reflect two criteria–goodness of fit and precision of estimation. Zellner (1992) noted that least squares estimators reflect goodness of fit consideration, while quadratic losses (which includes the squared error loss) are geared solely towards precision of estimation. He introduced the balanced loss to seek a trade-off between the two.

In Section 2 of this chapter, we introduce the balanced loss and derive the CB estimators under such a loss. Also this section develops the CB estimators for the one parameter natural exponential family of distributions with quadratic variance functions (NEF-QVF) as introduced in Morris (1982, 1983). We also provide an asymptotic expansion of the Bayes risks of the CB estimators in Section 2. Section 3 develops EB estimators under balanced loss functions and provides an asymptotic expansion of the MSE of such estimators that is correct up to $O(m^{-1})$. Second order correct estimators of the MSE's of these estimators are also provided. CEB estimators under balanced loss functions are derived in Section 4, and once again asymptotic expansions of their MSE's valid up to $O(m^{-1})$ are provided.

3.2 Constrained Bayes Estimators and Their Bayes Risks

Let $\boldsymbol{X} = (X_1, ..., X_m)^T$ with $E(\boldsymbol{X}) = \boldsymbol{\theta} = (\theta_1, ...\theta_m)^T$. For any estimator $\boldsymbol{e} = (e_1, ..., e_m)^T$ of $\boldsymbol{\theta}$, the balanced loss as introduced by Zellner (1988, 1992) is

$$L(\boldsymbol{\theta}, \boldsymbol{e}) = m^{-1}[w||\boldsymbol{X} - \boldsymbol{e}(\boldsymbol{X})||^2 + (1-w)||\boldsymbol{e}(\boldsymbol{X}) - \boldsymbol{\theta}||^2], \tag{3.1}$$

where $||\cdot||$ is the Euclidean norm and w $(0 \le w \le 1)$ is the known weight. The choice of w reflects the relative weight which the experimenter wants to assign to goodness of fit and precision of estimation. The extreme cases $w = 1$ and $w = 0$ refer solely to the precision of an estimate and goodness of fit respectively.

3.2.1 Constrained Bayes Estimators

Suppose now $\boldsymbol{e}^B(\boldsymbol{X})$ is a Bayes estimator of $\boldsymbol{\theta}$ under a prior π. Writing $\boldsymbol{e}^{PM}(\boldsymbol{x}) = E(\boldsymbol{\theta}|\boldsymbol{X} = \boldsymbol{x})$ and noting that $E[||\boldsymbol{e}(\boldsymbol{X}) - \boldsymbol{\theta}||^2|\boldsymbol{X} = \boldsymbol{x}] = E[tr\{V(\boldsymbol{\theta}|\boldsymbol{x})\} + ||\boldsymbol{e}(\boldsymbol{x}) - \boldsymbol{e}^{PM}(\boldsymbol{x})||^2]$, minimization of $E[L(\boldsymbol{\theta}, \boldsymbol{e})|\boldsymbol{X} = \boldsymbol{x}]$ with respect to $\boldsymbol{e}$ amounts to minimization of $w||\boldsymbol{x} - \boldsymbol{e}(\boldsymbol{x})||^2 + (1-w)||\boldsymbol{e}(\boldsymbol{x}) - \boldsymbol{e}^{PM}(\boldsymbol{x})||^2$ with respect to $\boldsymbol{e}$. A slight algebra shows that the minimizer $\boldsymbol{e}$ is given by

$$\boldsymbol{e}^B(\boldsymbol{x}) = w\boldsymbol{x} + (1-w)\boldsymbol{e}^{PM}(\boldsymbol{x}). \tag{3.2}$$

The estimate $\boldsymbol{e}^B$ is given in Zellner (1988, 1992).

However, the above estimate does not work well if one is interested in finding an optimal estimate of the histogram of the population parameters, namely, $m^{-1}\sum_{i=1}^m I_{[\theta_i \le t]}$. Indeed writing $\boldsymbol{e}^B(\boldsymbol{x}) = (e_1^B(\boldsymbol{x}), ..., e_m^B(\boldsymbol{x}))^T$, it is now clear that $\bar{e}^B(\boldsymbol{x}) = m^{-1}\sum_{i=1}^m e_i^B(\boldsymbol{x}) = w\bar{x} + (1-w)\bar{e}^{PM}(\boldsymbol{x}) \ne \bar{e}^{PM}(\boldsymbol{x})$ unless $w = 0$, where $\bar{e}^{PM}(\boldsymbol{x}) = m^{-1}\sum_{i=1}^m E(\theta_i|\boldsymbol{x})$. Also it is easy to check that $E[m^{-1}\sum_{i=1}^m(\theta_i - \bar{\theta})^2|\boldsymbol{x}] \ne m^{-1}\sum_{i=1}^m(e_i^B(\boldsymbol{x}) - \bar{e}^B(\boldsymbol{x}))^2$.

Following Louis (1984) and Ghosh (1992), we now seek compromise estimators $\boldsymbol{t} = (t_1, ..., t_m)^T$ of $\boldsymbol{\theta}$ which satisfy

$$\begin{aligned}
&(i) \;\; \bar{t} = m^{-1} \sum_{i=1}^{m} t_i = \bar{e}^{PM}(\boldsymbol{x}) \\
&(ii) \;\; E[m^{-1} \sum_{i=1}^{m} (t_i - \bar{t})^2 | \boldsymbol{x}] = m^{-1} \sum_{i=1}^{m} E[(\theta_i - \bar{\theta})^2 | \boldsymbol{x}],
\end{aligned} \tag{3.3}$$

and minimize $E[L(\boldsymbol{\theta}, \boldsymbol{t})|\boldsymbol{x}]$ with respect to $\boldsymbol{t}$ subject to (i) and (ii) in (3.3). The following theorem provides such compromise estimators.

A few notations are needed before stating the result. Let $H_1(\boldsymbol{x}) = \sum_{i=1}^{m} V(\theta_i - \bar{\theta}|\boldsymbol{x})$ and $H_2(\boldsymbol{x}) = \sum_{i=1}^{m}(e_i^{PM}(\boldsymbol{x}) - \bar{e}^{PM}(\boldsymbol{x}))^2$ so that $\sum_{i=1}^{m} E[(\theta_i - \bar{\theta})^2|\boldsymbol{x}] = H_1(\boldsymbol{x}) + H_2(\boldsymbol{x})$. Also let $b_i(\boldsymbol{x}) = w(x_i - \bar{x}) + (1-w)(e_i^{PM} - \bar{e}^{PM})$, $1 \le i \le m$. Then the following result holds.

<u>Theorem 3.1</u> Assume (3.3). Then $E[L(\boldsymbol{\theta}, \boldsymbol{t})|\boldsymbol{x}]$ is minimized with respect to $\boldsymbol{t}$ when

$$t_i \equiv t_i(\boldsymbol{x}) = \left[\frac{H_1(\boldsymbol{x}) + H_2(\boldsymbol{x})}{\sum_{i=1}^{m} b_i^2(\boldsymbol{x})} \right]^{1/2} b_i(\boldsymbol{x}) + \bar{e}^{PM}(\boldsymbol{x}), \;\; i = 1, ..., m. \tag{3.4}$$

<u>Proof of Theorem 3.1</u> Our objective is to minimize $\sum_{i=1}^{m} E[w(t_i - x_i)^2 + (1-w)(t_i - \theta_i)^2|\boldsymbol{x}]$ subject (3.3). Since $E[(t_i - \theta_i)^2|\boldsymbol{x}] = V(\theta_i|\boldsymbol{x}) + (t_i - e_i^{PM}(\boldsymbol{x}))^2$, this amounts to minimization of $\sum_{i=1}^{m}[w(t_i - x_i)^2 + (1-w)(t_i - e_i^{PM}(\boldsymbol{x}))^2]$ with respect to $\boldsymbol{t}$ subject to (3.3). Let

$$\begin{aligned}
g(\boldsymbol{t}) \;\; = \;\; & \sum_{i=1}^{m}[w(t_i - x_i)^2 + (1-w)(t_i - e_i^{PM}(\boldsymbol{x}))^2 \\
& -2\lambda_1(\bar{t} - \bar{e}^{PM}(\boldsymbol{x})) - 2\lambda_2\{\sum_{i=1}^{m}(t_i - \bar{t})^2 - (H_1(\boldsymbol{x}) + H_2(\boldsymbol{x}))\}],
\end{aligned}$$

where λ_1 and λ_2 are Lagrange multipliers. Differentiation with respect to t_i gives

$$\begin{aligned}
0 = \frac{\partial g}{\partial t_i} \;\; = \;\; & 2w(t_i - x_i) + 2(1-w)(t_i - e_i^{PM}(\boldsymbol{x})) \\
- \;\; & 2\lambda_1 - 2\lambda_2(t_i - \bar{t})(1 - m^{-1}).
\end{aligned}$$

Summing over i, by (i) of (3.3), $0 = 2w(\bar{t} - \bar{x}) - 2\lambda_1$, i.e., $\lambda_1 = w(\bar{t} - \bar{x})$. Next writing $\lambda_2' = \lambda_2(1 - m^{-1})$, again by (i) of (3.3),

$$(1 - \lambda_2')(t_i - \bar{t}) = w(x_i - \bar{x}) + (1 - w)(e_i^{PM}(\boldsymbol{x}) - \bar{e}^{PM}(\boldsymbol{x})) = b_i, \ 1 \le i \le m. \quad (3.5)$$

This implies $(1 - \lambda_2')^2 \sum_{i=1}^m (t_i - \bar{t})^2 = \sum_{i=1}^m b_i^2$. Then by (ii) of (3.3),

$$(1 - \lambda_2')^2 [H_1(\boldsymbol{x}) + H_2(\boldsymbol{x})] = \sum_{i=1}^m b_i^2. \quad (3.6)$$

Now from (3.5) and (3.6), $t_i - \bar{t} = [H_1(\boldsymbol{x}) + H_2(\boldsymbol{x})]^{1/2} (\sum_{i=1}^m b_i^2)^{-1/2} b_i$, i.e.,

$$t_i = [H_1(\boldsymbol{x}) + H_2(\boldsymbol{x})]^{1/2} b_i / (\sum_{i=1}^m b_i^2)^{-1/2} + \bar{e}^{PM}(\boldsymbol{x}),$$

again by (i) of (3.3).

Next we illustrates an application of the above result for the one-parameter exponential family.

3.2.2 One-Parameter Exponential Family

Let X_i denote the sample average of a random sample of size n from the one-parameter exponential family with mean $\psi'(\phi_i)$. Thus X_i has pdf

$$f(x_i | \phi_i) = exp[n(\phi_i x_i - \psi(\phi_i)) + c(x_i)], \ 1 \le i \le m.$$

Consider the conjugate prior for ϕ_i given by

$$\pi(\phi_i | m, \lambda) = exp[\lambda(\phi_i m - \psi(\phi_i)) + g(m, \lambda)].$$

Then the population mean $\psi'(\phi_i)$ has posterior expectation

$$E[\psi'(\phi_i) | x_i] = (1 - B)x_i + Bm,$$

where $B = \frac{\psi''(\theta_i)/n}{\psi''(\theta_i)/n + \psi''(\theta_i)/\lambda} = \frac{\lambda}{n + \lambda}$. In this case $e_i^{PM}(\boldsymbol{x}) - \bar{e}^{PM}(\boldsymbol{x}) = (1 - B)(x_i - \bar{x})$, $\bar{x} = m^{-1} \sum_{i=1}^m x_i$, and

$$\begin{aligned} b_i(x) &= w(x_i - \bar{x}) + (1-w)(e_i^{PM}(\boldsymbol{x}) - \bar{e}^{PM}(\boldsymbol{x})) \\ &= \{1-(1-w)B\}(x_i - \bar{x}). \end{aligned}$$

Also,

$$H_1(\boldsymbol{x}) = \sum_{i=1}^{m} V(\theta_i - \bar{\theta}|\boldsymbol{x}) = (n+\lambda)^{-1}\sum_{i=1}^{m} E[\psi''(\phi_i)|x_i](1-m^{-1}),$$

and

$$H_2(\boldsymbol{x}) = \sum_{i=1}^{m}(e_i^{PM}(\boldsymbol{x}) - \bar{e}^{PM}(\boldsymbol{x}))^2 = (1-B)^2\sum_{i=1}^{m}(x_i - \bar{x})^2.$$

Now writing $a(\boldsymbol{x}) = [1 + H_1(\boldsymbol{x})/H_2(\boldsymbol{x})]^{1/2}$, the CB estimator of θ_i is given by

$$\begin{aligned} \hat{\theta}_i^{CB} &= a(\boldsymbol{x})\frac{(1-B)\{\sum_{i=1}^{m}(x_i-\bar{x})^2\}^{1/2}}{\{1-(1-w)B\}^2\sum_{i=1}^{m}(x_i-\bar{x})^2\}^{1/2}}(1-(1-w)B)(x_i-\bar{x}) + \bar{e}^{PM}(\boldsymbol{x}) \\ &= a(\boldsymbol{x})(e_i^{PM}(\boldsymbol{x}) - \bar{e}^{PM}(\boldsymbol{x})) + \bar{e}^{PM}(\boldsymbol{x}) \\ &= a(\boldsymbol{x})e_i^{PM}(\boldsymbol{x}) + (1-a(\boldsymbol{x}))\bar{e}^{PM}(\boldsymbol{x}). \end{aligned}$$

This is same as the expression obtained by Ghosh (1992) assuming squared error loss. In a way, this demonstrates the loss robustness of the estimator given in (3.4).

For the special case of the natural exponential family with quadratic variance functions (NEF-QVF) as proposed by Morris (1982,1983), $\psi''(\phi_i) = \nu_0 + \nu_1\psi'(\phi_i) + \nu_2(\psi'(\phi_i))^2$, where ν_0, ν_1 and ν_2 are not simultaneously zeros. Then,

$$\begin{aligned} H_1(\boldsymbol{x}) &= (n+\lambda-\nu_2)^{-1}(1-m^{-1}) \times \\ &\quad [m\nu_0 + m\nu_1\{(1-B)\bar{x} + B\mu\} + \nu_2\sum_{i=1}^{m}\{(1-B)x_i + B\mu\}^2] \\ &= (n+\lambda-\nu_2)^{-1}(m-1) \times \\ &\quad [\nu_0 + \nu_1((1-B)\bar{x} + B\mu) + \nu_2((1-B)\bar{x} + B\mu)^2 + \nu_2(1-B)^2\sum_{i=1}^{m}(x_i-\bar{x})^2]. \end{aligned}$$

3.2.3 Bayes Risk of the Constrained Bayes Estimators

We now consider the normal-normal example when $X_i|\theta_i$ are independent $N(\theta_i, 1)$ and θ_i are *iid* $N(\mu, \tau^2)$. In this case $\psi(\theta_i) = \frac{1}{2}\theta_i^2$ so that $\phi_i = \psi'(\theta_i) = \theta_i$ and $\psi''(\theta_i) = 1$. Also, $B = (1+\tau^2)^{-1}$. Now $H_1(\boldsymbol{X}) = (m-1)(1-B)$ and $a(\boldsymbol{X}) = [1+(1-B)^{-1}S^{-1}]^{1/2}$, where $S = \sum_{i=1}^m (X_i - \bar{X})^2/(m-1)$. The posterior mean of $\boldsymbol{\theta}$ is $\boldsymbol{e}^{PM}(\boldsymbol{X}) = (1-B)\boldsymbol{X} + B\mu\boldsymbol{1}_m$ with Bayes risk $1-B$. We now find the Bayes risk of $\hat{\boldsymbol{\theta}}^{CB}$ in the following theorem.

Theorem 3.2 Under the loss given in (3.1), the Bayes risk of $\hat{\boldsymbol{\theta}}^{CB}$ is given by

$$w + a_1(B) + \frac{1}{m}[(1-2w)(1-B) - a_1(B) + a_2(B)] + O(m^{-3/2}),$$

where $a_1(B) = 2B^{-1}(1-B) - 2B^{-1}(1-B)^{1/2}\{1-(1-w)B\}$, and $a_2(B) = \frac{B}{2}(1-B)^{1/2}\{1-(1-w)B\}$.

Proof of Theorem 3.2 First we can express $\hat{\boldsymbol{\theta}}^{CB}$ as follows.

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{CB} &= a(\boldsymbol{X})[(1-B)\boldsymbol{X} + B\mu\boldsymbol{1}_m] + (1-a(\boldsymbol{X}))[(1-B)\bar{X} + B\mu]\boldsymbol{1}_m \\ &= a(\boldsymbol{X})(1-B)(\boldsymbol{X} - \bar{X}\boldsymbol{1}_m) + [(1-B)\bar{X} + B\mu]\boldsymbol{1}_m. \end{aligned}$$

And then, the Bayes risk under balanced loss functions is

$$E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{CB})] = m^{-1}E[w||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CB}||^2 + (1-w)E||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{CB}||^2]. \tag{3.7}$$

But

$$\begin{aligned} \boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CB} &= \boldsymbol{X} - (1-B)a(\boldsymbol{X})(\boldsymbol{X} - \bar{X}\boldsymbol{1}_m) - [(1-B)\bar{X} + B\mu]\boldsymbol{1}_m \\ &= [1-(1-B)a(\boldsymbol{X})](\boldsymbol{X} - \bar{X}\boldsymbol{1}_m) + B(\bar{X} - \mu)\boldsymbol{1}_m. \end{aligned}$$

By the independence of $\bar{X}$ and $\sum_{i=1}^m (X_i - \bar{X})^2$, one gets

$$m^{-1}E||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CB}||^2 = m^{-1}E[\{1-(1-B)a(\boldsymbol{X})\}^2(m-1)S] + Bm^{-1}. \tag{3.8}$$

Again,

$$\begin{aligned} m^{-1}E||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{CB}||^2 &= m^{-1}E||\boldsymbol{\theta} - \boldsymbol{e}^{PM}(\boldsymbol{X}) + \boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{CB}||^2 \\ &= 1 - B + m^{-1}E||\boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{CB}||^2. \end{aligned} \tag{3.9}$$

But

$$\begin{aligned} \boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{CB} &= (1-B)\boldsymbol{X} + B\mu\mathbf{1}_m - a(\boldsymbol{X})[(1-B)\boldsymbol{X} + B\mu\mathbf{1}_m] \\ &- (1-a(\boldsymbol{X}))[(1-B)\bar{X} + B\mu]\mathbf{1}_m \\ &= (1-a(\boldsymbol{X}))(1-B)(\boldsymbol{X} - \bar{X}\mathbf{1}_m). \end{aligned}$$

Hence,

$$m^{-1}E||\boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{CB}||^2 = (1-B)^2\{(m-1)/m\}E[(a(\boldsymbol{X})-1)^2S]. \tag{3.10}$$

Combining (3.7)-(3.10), we get

$$\begin{aligned} E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{CB})] &= (1-w)(1-B) + wBm^{-1} \\ &+ \frac{m-1}{m}E[w\{1-(1-B)a(\boldsymbol{X})\}^2S] \\ &+ \frac{m-1}{m}E[(1-w)(1-B)^2(a(\boldsymbol{X})-1)^2S]. \end{aligned}$$

Next we simplify

$$\begin{aligned} & w[1-(1-B)a(\boldsymbol{X})]^2 + (1-w)(1-B)^2(a(\boldsymbol{X})-1)^2 \\ & = (1-B)^2a^2(\boldsymbol{X}) + w + (1-w)(1-B)^2 \\ & -2a(\boldsymbol{X})(1-B)[1-(1-w)B]. \end{aligned} \tag{3.11}$$

Now, $E[a^2(\boldsymbol{X})S] = E[\{1+(1-B)^{-1}S^{-1}\}S] = B^{-1}+(1-B)^{-1} = B^{-1}(1-B)^{-1}$.
Hence, from (3.11),

$$E[\{w(1-(1-B)a(\boldsymbol{X}))^2+(1-w)(1-B)^2(a(\boldsymbol{X})-1)^2\}S]$$
$$= B^{-1}(1-B)+B^{-1}[w+(1-w)(1-B)^2]-2(1-B)[1-(1-w)B]E[a(\boldsymbol{X})S]$$
$$= B^{-1}(1-B)+B^{-1}-2(1-w)+(1-w)B$$
$$-2(1-B)[1-(1-w)B]E[a(\boldsymbol{X})S]. \tag{3.12}$$

Next

$$E[a(\boldsymbol{X})S] = (1-B)^{-1/2}E[\{(1-B)S^2+S\}^{1/2}] = (1-B)^{-1/2}E[g(S)] \quad (say). \tag{3.13}$$

By the Taylor expansion,

$$\begin{aligned} g(S) &= g(ES)+(S-ES)g'(ES)+\frac{1}{2}(S-ES)^2g''(ES) \\ &+ \frac{1}{2}(S-ES)^3\int_0^1(1-\lambda)^2g'''[\lambda S+(1-\lambda)ES]d\lambda. \end{aligned} \tag{3.14}$$

Note that

$$\begin{aligned} g(x) &= [x+(1-B)x^2]^{1/2}, \\ g'(x) &= \frac{1+2(1-B)x}{2[x+(1-B)x^2]^{1/2}}, \\ g''(x) &= \frac{1-B}{x+(1-B)x^2}-\frac{[1+2(1-B)x]^2}{4[x+(1-B)x^2]^{3/2}} = -\frac{1}{4[x+(1-B)x^2]^{3/2}}, \\ g'''(x) &= \frac{3[1+2(1-B)x]}{8[x+(1-B)x^2]^{5/2}}. \end{aligned}$$

Noting that $E(S) = B^{-1}$,

$$g(ES) = [B^{-1}+(1-B)B^{-2}]^{1/2} = B^{-1}. \tag{3.15}$$

$$g''(ES) = -\frac{1}{4[B^{-1}+(1-B)B^{-2}]^{3/2}} = -\frac{B^3}{4}. \tag{3.16}$$

Finally, since $x > 0$, $|g'''(x)| \le \frac{3(1+2x)}{8x^{5/2}(1-B)^{5/2}}$. Thus,

$$\begin{aligned} |g'''[\lambda S + (1-\lambda)ES]| &\le (1-B)^{-5/2} \times \\ &\quad \left[\frac{3}{8}[\lambda S + (1-\lambda)ES]^{-5/2} + \frac{3}{4}[\lambda S + (1-\lambda)ES]^{-3/2}\right] \\ &\le (1-B)^{-5/2} \times \\ &\quad \left[\frac{3}{8}(1-\lambda)^{-5/2}B^{5/2} + \frac{3}{4}(1-\lambda)^{-3/2}B^{3/2}\right]. \end{aligned} \tag{3.17}$$

Hence, from (3.17), $\int_0^1 (1-\lambda)^2 |g'''[\lambda S + (1-\lambda)ES]| d\lambda < \infty$.

Also $E(S - ES)^2 = 2B^{-2}/(m-1)$, $E|S - ES|^3 = O(m^{-3/2})$. Hence combining (3.15)-(3.17), one gets

$$\begin{aligned} E[g(S)] &= B^{-1} + \frac{B^{-2}}{m-1}(-\frac{B^3}{4}) + O(m^{-3/2}) \\ &= B^{-1} - \frac{B}{4m} + O(m^{-3/2}). \end{aligned} \tag{3.18}$$

Thus, from (3.12)-(3.14) and (3.18), one gets the result.

3.3 Empirical Bayes Estimators and Their Bayes Risks

In this section we discuss the empirical Bayes estimators under balanced loss function and also the Bayes risk of the such estimators valid up to $O(m^{-1})$ given loss functions.

3.3.1 Empirical Bayes Estimators

We continue with the normal-normal scenario where as before $X_i|\theta_i \sim^{ind} N(\theta_i, 1)$ and $\theta_i \sim^{iid} N(\mu, \tau^2)$. However, this time μ and τ^2 are both unknown and need to be estimated from the marginal distributions of the X_i's. Marginally, X_i are *iid* $N(\mu, B^{-1})$, where $B = (1+\tau^2)^{-1}$. Based on these marginals, $\bar{X} = m^{-1}\sum_{i=1}^m X_i$, and $S = \sum_{i=1}^m (X_i - \bar{X})^2/(m-1)$ is complete sufficient for (μ, B). Following Morris (1981), we estimate μ and B respectively by $\hat{\mu} = \bar{X}$ and $\hat{B} = min(\frac{m-3}{m-1}, \frac{m-3}{(m-1)S})$. We now have the EB estimator of $\boldsymbol{\theta}$ from (3.2) as

$$\begin{aligned}\hat{\boldsymbol{\theta}}^{EB} &= w\boldsymbol{X} + (1-w)[(1-\hat{B})\boldsymbol{X} + \hat{B}\bar{X}\mathbf{1}_m] \\ &= [1-(1-w)\hat{B}]\boldsymbol{X} + (1-w)\hat{B}\bar{X}\mathbf{1}_m.\end{aligned}$$

3.3.2 Bayes Risk of the Empirical Bayes Estimators

We now find the the Bayes risk of the EB estimator correct to $O(m^{-1})$ under the loss (3.1).

Thorem 3.3 Under the loss given in (3.1), the Bayes risk of $\hat{\boldsymbol{\theta}}^{EB}$ is given by

$$(1-w)(1-(1-w)B) + \frac{3(1-w)^2 B}{m} + o(m^{-1}).$$

Proof of Theorem 3.3 To this end, first let $\hat{\hat{B}} = \frac{m-3}{(m-1)S}$. We begin with

$$\begin{aligned}\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{EB} &= \boldsymbol{X} - [1-(1-w)\hat{B}]\boldsymbol{X} - (1-w)\hat{B}\bar{X}\mathbf{1} \\ &= (1-w)\hat{B}(\boldsymbol{X} - \bar{X}\mathbf{1}) \\ &= (1-w)[\hat{\hat{B}}(\boldsymbol{X} - \bar{X}\mathbf{1}) + (\hat{B} - \hat{\hat{B}})(\boldsymbol{X} - \bar{X}\mathbf{1})].\end{aligned}$$

Hence,

$$\begin{aligned}E[||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{EB}||^2] &= (1-w)^2 E[\hat{\hat{B}}^2(m-1)S] + (1-w)^2 E[(\hat{B} - \hat{\hat{B}})^2(m-1)S] \\ &+ 2(1-w)^2 E[\hat{\hat{B}}(\hat{B} - \hat{\hat{B}})(m-1)S]; \qquad (3.19)\end{aligned}$$

$$E[\hat{\hat{B}}^2(m-1)S] = E[\frac{(m-3)^2}{(m-1)S}] = (m-3)B; \qquad (3.20)$$

$$\begin{aligned}E[(\hat{B} - \hat{\hat{B}})^2(m-1)S] &= E[(\frac{m-3}{m-1} - \hat{\hat{B}})^2(m-1)S I_{[\frac{m-3}{(m-1)S} > \frac{m-3}{m-1}]}] \\ &\leq E^{1/2}[(\frac{m-3}{m-1} - \hat{\hat{B}})^4(m-1)^2S^2] \times \\ &\quad P^{1/2}[\frac{1}{S} > 1]. \qquad (3.21)\end{aligned}$$

But by the c_δ-inequalty $(a+b)^{1+\delta} \le 2^\delta(a^{1+\delta}+b^{1+\delta})$ for $a>0$, $b>0$ and $\delta>0$,

$$\begin{aligned} E[(\frac{m-3}{m-1}-\hat{B})^4(m-1)^2S^2] &\le 8E[(\frac{m-3}{m-1}+\hat{B}^4)(m-1)^2S^2] \\ &= 8[\frac{(m-3)^4(m+1)B^{-2}}{(m-1)^3}+\frac{(m-3)^4B^2}{(m-3)(m-5)}] \\ &= O(m^2), \end{aligned} \tag{3.22}$$

while,

$$\begin{aligned} P[\frac{m-3}{(m-1)S}>1] &= P[S<1] \\ &= P(B(m-1)S<B(m-1)) \\ &= P(\chi^2_{m-1}-(m-1)<(m-1)(B-1)) \\ &\le P(|\chi^2_{m-1}-(m-1)|>(m-1)(1-B)) \\ &= O(m^{-r}), \end{aligned} \tag{3.23}$$

for any arbitrary $r>0$. Choosing $r>4$, one gets from (3.21)-(3.23),

$$E[(\hat{B}-\hat{\hat{B}})^2(m-1)S]=o(m^{-1}). \tag{3.24}$$

Next by the Schwarz inequality, (3.20) and (3.24),

$$\begin{aligned} E[\hat{\hat{B}}(\hat{B}-\hat{\hat{B}})(m-1)S] &\le E^{1/2}[\hat{\hat{B}}^2(m-1)S]E^{1/2}[(\hat{B}-\hat{\hat{B}})^2(m-1)S] \\ &= O(m^{1/2-r/2})=o(1) \quad for\ r>2. \end{aligned} \tag{3.25}$$

Hence, from (3.19), (3.20), (3.24) and (3.25),

$$m^{-1}E[||\boldsymbol{X}-\hat{\boldsymbol{\theta}}^{EB}||^2]=\frac{(m-3)(1-w)^2B}{m}+o(m^{-1}). \tag{3.26}$$

Next we calculate

$$\begin{aligned} E[||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{EB}||^2] &= E||\boldsymbol{\theta} - \boldsymbol{e}^{PM}(\boldsymbol{X}) + \boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{EB}||^2 \\ &= m(1-B) + E||\boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{EB}||^2. \end{aligned} \tag{3.27}$$

Also,

$$\begin{aligned} \boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{EB} &= (1-B)\boldsymbol{X} + B\mu\mathbf{1}_m - [1-(1-w)\hat{B}]\boldsymbol{X} - (1-w)\hat{B}\bar{X}\mathbf{1}_m \\ &= [(1-w)\hat{B} - B]\boldsymbol{X} - [(1-w)\hat{B} - B\}\bar{X}\mathbf{1}_m - B(\bar{X} - \mu)\mathbf{1}_m \\ &= [(1-w)\hat{B} - B](\boldsymbol{X} - \bar{X}\mathbf{1}_m) - B(\bar{X} - \mu)\mathbf{1}_m. \end{aligned}$$

Hence, by the independence of $\bar{X}$ and $\boldsymbol{X} - \bar{X}\mathbf{1}_m$,

$$E||\boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{EB}||^2 = E[\{(1-w)\hat{B} - B\}^2(m-1)S] + B \tag{3.28}$$

Also,

$$\begin{aligned} E[\{(1-w)\hat{B} - B\}^2(m-1)S] &= E[\{(1-w)\hat{\hat{B}} - B\}^2(m-1)S] \\ &+ (1-w)^2 E[(\hat{B} - \hat{\hat{B}})^2(m-1)S] \\ &+ 2(1-w) \times \\ &\quad E[((1-w)\hat{\hat{B}} - B)(\hat{B} - \hat{\hat{B}})(m-1)S]. \end{aligned} \tag{3.29}$$

Next

$$\begin{aligned} E[\{(1-w)\hat{\hat{B}} - B\}^2(m-1)S] &= (1-w)^2 E[\hat{\hat{B}}^2(m-1)S] + E[B^2(m-1)S] \\ &- 2(1-w)BE[\hat{\hat{B}}(m-1)S] \\ &= (1-w)^2 E[\frac{(m-3)^2}{(m-1)S}] \\ &- 2(1-w)(m-3)B + (m-1)B \\ &= (1-w)^2(m-3)B \\ &- 2(1-w)(m-3)B + (m-1)B \end{aligned}$$

$$
\begin{aligned}
&= (m-1)B - (1-w^2)(m-3)B \\
&= [2+(m-3)w^2]B. \qquad (3.30)
\end{aligned}
$$

By (3.24), (3.30) and the Schwarz inequality, for any $r > 3/2$,

$$
\begin{aligned}
E[|(1-w)\hat{B} - B||\hat{B} - \hat{\hat{B}}|(m-1)S] &\leq E^{1/2}[\{(1-w)\hat{\hat{B}} - B\}^2(m-1)S] \times \\
&\quad E^{1/2}[(\hat{B} - \hat{\hat{B}})^2(m-1)S] \\
&= [(2+(m-3)w^2)B]^{1/2})O(m^{-r}) \\
&= O(m^{1/2})o(m^{-1}) = o(1). \qquad (3.31)
\end{aligned}
$$

Now by (3.28), (3.29), (3.31) and (3.24), one gets

$$
m^{-1}E||e^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{EB}||^2 = \frac{B}{m} + \frac{(2+(m-3)w^2)B}{m} + o(m^{-1}). \qquad (3.32)
$$

Combining (3.26), (3.27) and (3.32),

$$
\begin{aligned}
E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{EB})] &= w(1-w)^2[B - \frac{3B}{m} + o(m^{-1})] \\
&+ (1-w)[1 - B + \frac{B}{m} + \frac{(2+(m-3)w^2)B}{m} + o(m^{-1})] \\
&= (1-w)(1-(1-w)B) + \frac{3(1-w)^2B}{m} + o(m^{-1}). \qquad (3.33)
\end{aligned}
$$

Thus, from (3.33), one gets the result.

To estimate the MSE expression given in (3.33), we need only to find $E(\hat{B})$. With the same argument as before

$$
\begin{aligned}
E[\hat{B}] &= E[\hat{B} - \hat{\hat{B}}] + E[\hat{\hat{B}}] \\
&= E[(\frac{m-3}{m-1} - \hat{\hat{B}})I_{[\frac{1}{S}>1]}] + E[\frac{m-3}{(m-1)S}]; \qquad (3.34)
\end{aligned}
$$

$$
E[\frac{m-3}{(m-1)S}] = B; \qquad (3.35)
$$

$$E[(\frac{m-3}{m-1} - \hat{\hat{B}})I_{[\frac{1}{S}>1]}] \leq E^{1/2}[(\frac{m-3}{m-1} - \hat{\hat{B}})^2]P^{1/2}(S < 1). \tag{3.36}$$

But,

$$E[(\frac{m-3}{m-1} - \hat{\hat{B}})^2] \leq 4E[\frac{m-3}{m-1} + \hat{\hat{B}}^2] = 4(\frac{m-3}{m-1} + \frac{(m-3)B^2}{m-5}) = O(1). \tag{3.37}$$

Hence, from (3.36), (3.37) and (3.23),

$$E[\hat{B} - \hat{\hat{B}}] = o(m^{-1}). \tag{3.38}$$

Now by (3.34), (3.35) and (3.38), one gets

$$E[\hat{B}] = B + o(m^{-1}).$$

Hence,

$$\begin{aligned} E[(1-w)(1-(1-w)\hat{B}) + \frac{3(1-w)^2\hat{B}}{m}] &= [(1-w)(1-(1-w)B) + \frac{3(1-w)^2 B}{m} \\ &+ o(m^{-1}). \end{aligned}$$

3.4 Constrained Empirical Bayes Estimators

Constrained EB estimators are obtained by substituting $\hat{B}$ for B and $\bar{X}$ for μ in constrained Bayes estimators. Accordingly,

$$\hat{\boldsymbol{\theta}}^{CEB} = a_{EB}[(1-\hat{B})\boldsymbol{X} + \hat{B}\bar{X}\mathbf{1}_m] + (1-a_{EB})\bar{X}\mathbf{1}_m,$$

where $a_{EB} = [1 + (1-\hat{B})^{-1}S^{-1}]^{1/2}$. Now under the loss (3.1), the the Bayes risk of the constrained EB estimator is

$$E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{CEB})] = m^{-1}[wE||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}||^2 + (1-w)E||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{CEB}||^2].$$

We now find the the Bayes risk of the constrained EB estimator correct to $O(m^{-1})$ under the loss (3.1).

Thorem 3.4 Under the loss given in (3.1), the Bayes risk of $\hat{\boldsymbol{\theta}}^{CEB}$ is given by

$$w + c_1(B) + \frac{1}{m}[3 - 2wB - c_1(B) + 2(1-(1-w)B)c_2(B)] + O(m^{-3/2}),$$

where $c_1(B) = 2[B^1(1 - B - (1-B)^{1/2}) + (1-w)(1-B)^{1/2}$ and $c_2(B) = 1 - (2-B)(1-B)^{-1/2} + \frac{1}{4}B(1-B)^{-3/2}$.

Proof of Theorem 3.4 First we write

$$\begin{aligned} ||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}||^2 &= ||[1 - a_{EB}(1-\hat{B})](\boldsymbol{X} - \bar{X}\mathbf{1}_m)||^2 \\ &= [1 - a_{EB}(1-\hat{B})]^2(m-1)S \\ &= [1 + a_{EB}^2(1-\hat{B})^2 - 2a_{EB}(1-\hat{B})](m-1)S \\ &= [1 + (1-\hat{B})^2 + (1-\hat{B})S^{-1} - 2a_{EB}(1-\hat{B})](m-1)S. \end{aligned}$$

Accordingly,

$$\begin{aligned} E||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}||^2 &= (m-1)B^{-1} + E[(1-\hat{B})^2(m-1)S] + E[(1-\hat{B})(m-1)] \\ &- 2E[(1-\hat{B})a_{EB}(m-1)S] \end{aligned}$$

Noting that $\hat{\hat{B}} = \frac{m-3}{(m-1)S}$,

$$\begin{aligned} E[(1-\hat{B})^2(m-1)S] &= E[(1-\hat{\hat{B}} + \hat{\hat{B}} - \hat{B})^2(m-1)S] \\ &= E[(1-\hat{\hat{B}})^2(m-1)S] + E[(\hat{\hat{B}} - \hat{B})^2(m-1)S] \\ &+ 2E[(1-\hat{\hat{B}})(\hat{\hat{B}} - \hat{B})(m-1)S]. \end{aligned} \tag{3.39}$$

We calculate

$$\begin{aligned} E[(1-\hat{\hat{B}})^2(m-1)S] &= B^{-1}(m-1) - 2(m-3) + (m-3)B \\ &= 2B^{-1} + (m-3)B^{-1}(1-B)^2; \end{aligned} \tag{3.40}$$

Now by (3.20), (3.21), (3.40) and Schwarz inequality for the rightmost term of (3.39), one gets

$$E[(1-\hat{B})^2(m-1)S] = 2B^{-1} + (m-3)B^{-1}(1-B)^2 + o(m^{-1}). \tag{3.41}$$

Also,

$$\begin{aligned} E[(1-\hat{B})(m-1)] &= E[(1-\hat{\hat{B}})(m-1) + (\hat{\hat{B}} - \hat{B})(m-1)] \\ &= (m-1)(1-B) + o(m^{-1}). \end{aligned} \tag{3.42}$$

Thus, from (3.40) and (3.42)

$$\begin{aligned} E||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}||^2 &= (m-1)B^{-1} + 2B^{-1} + (m-3)B^{-1}(1-B)^2 + (m-1)(1-B) \\ &- 2E[(1-\hat{B})a_{EB}(m-1)S] + o(m^{-1}). \end{aligned} \tag{3.43}$$

Next we find

$$\begin{aligned} E||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{CEB}||^2 &= E||\boldsymbol{\theta} - \boldsymbol{e}^{PM}(\boldsymbol{X}) + \boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{CEB}||^2 \\ &= m(1-B) + E||\boldsymbol{e}^{PM}(\boldsymbol{X}) - \hat{\boldsymbol{\theta}}^{CEB}||^2. \end{aligned}$$

We now write

$$\begin{aligned} \hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{e}^{PM}(\boldsymbol{X}) &= a_{EB}[(1-\hat{B})\boldsymbol{X} + \hat{B}\bar{X}\boldsymbol{1}_m] + (1-a_{EB})\bar{X}\boldsymbol{1}_m \\ &- (1-B)\boldsymbol{X} - B\mu\boldsymbol{1}_m \\ &= [a_{EB}(1-\hat{B}) - (1-B)](\boldsymbol{X} - \bar{X}\boldsymbol{1}_m) + B(\bar{X} - \mu)\boldsymbol{1}_m. \end{aligned}$$

Once again, by the independence of $\boldsymbol{X} - \bar{X}\boldsymbol{1}_m$ and $\bar{X} - \mu$,

$$E||\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{e}^{PM}(\boldsymbol{X})||^2 = B + E[\{a_{EB}(1-\hat{B}) - (1-B)\}^2(m-1)S].$$

Now from (3.41), (3.42) and

$$\begin{aligned} &E[\{a_{EB}(1-\hat{B}) - (1-B)\}^2(m-1)S] \\ &= E[\{a_{EB}^2(1-\hat{B})^2 + (1-B)^2 - 2(1-B)a_{EB}(1-\hat{B})\}(m-1)S] \end{aligned}$$

$$= E[(1-\hat{B})^2(m-1)S] + E[(1-\hat{B})(m-1)] + E[(1-B)^2(m-1)S]$$
$$-2(1-B)E[a_{EB}(1-\hat{B})(m-1)S],$$

one gets

$$\begin{aligned}
E||\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{\theta}||^2 &= m(1-B) + B + 2B^{-1} + (m-3)B^{-1}(1-B)^2 \\
&+ (m-1)(1-B) + (m-1)(1-B)^2B^{-1} \\
&- 2(1-B)E[a_{EB}(1-\hat{B})(m-1)S] + o(m^{-1}). \qquad (3.44)
\end{aligned}$$

Now by (3.43) and (3.44), the MSE of CEB can be expressed as

$$\begin{aligned}
E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{CEB})] &= m^{-1}[wE||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}||^2 + (1-w)E||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}^{CEB}||^2] \\
&= w + 2B^{-1}(1-B) + \frac{1}{m}[3 - 2w - 2B^{-1}(1-B) + 2(1-B)] \\
&- \frac{2}{m}\{1-(1-w)B\}E[a_{EB}(1-\hat{B})(m-1)S] + o(m^{-1}). \qquad (3.45)
\end{aligned}$$

But

$$\begin{aligned}
E[a_{EB}(1-\hat{B})S] &= E[(1-\hat{B})^2S^2 + (1-\hat{B})S]^{1/2} \\
&= E[\{(1-\hat{B})^2S^2 + (1-\hat{B})S\}^{1/2}I_{[\hat{B}\neq\hat{\hat{B}}]}] \\
&+ E[\{(1-\hat{B})^2S^2 + (1-\hat{B})S\}^{1/2}I_{[\hat{B}=\hat{\hat{B}}]}] \\
&= E[\{(1-\frac{m-3}{m-1})^2S^2 + (1-\frac{m-3}{m-1})S\}^{1/2}I_{[S<1]}] \\
&+ E[\{(1-\hat{\hat{B}})^2S^2 + (1-\hat{\hat{B}})S\}^{1/2}I_{[S>1]}].
\end{aligned}$$

But

$$\begin{aligned}
E[\{(1-\hat{\hat{B}})^2S^2 + (1-\hat{\hat{B}})S\}^{1/2}] &= E[\{(1-\hat{\hat{B}})^2S^2 + (1-\hat{\hat{B}})S\}^{1/2}I_{[S>1]}] \\
&+ E[\{(1-\hat{\hat{B}})^2S^2 + (1-\hat{\hat{B}})S\}^{1/2}I_{[S<1]}] \\
&= E[\{(1-\hat{\hat{B}})^2S^2 + (1-\hat{\hat{B}})S\}^{1/2}I_{[S>1]}] \\
&+ o(m^{-1}).
\end{aligned}$$

From the above result,

$$E[(1-\hat{B})^2S^2+(1-\hat{B})S]^{1/2} = E[\{(1-\hat{\hat{B}})^2S^2+(1-\hat{\hat{B}})S\}^{1/2}]+o(m^{-1}). \quad (3.46)$$

Again,

$$\begin{aligned}
E[(1-\hat{\hat{B}})S\{1+(1-\hat{\hat{B}})S\}]^{1/2} &= E[(S-\frac{m-3}{m-1})(1+S-\frac{m-3}{m-1})]^{1/2} \\
&= E[(S-\frac{m-3}{m-1})(S+\frac{2}{m-1})]^{1/2} \\
&= E[S^2-\frac{m-5}{m-1}S-\frac{2(m-3)}{(m-1)^2}]^{1/2} \\
&= E[h(S)]. \quad (say)
\end{aligned}$$

Writing for $x > 1$, m large, $h(x) = [x^2 - \frac{m-5}{m-1}x - \frac{2(m-3)}{(m-1)^2}]^{1/2}$,

$$h'(x) = \frac{1}{2}[x^2-\frac{m-5}{m-1}x-\frac{2(m-3)}{(m-1)^2}]^{-1/2}(2x-\frac{m-5}{m-1});$$

$$\begin{aligned}
h''(x) &= -\frac{1}{4}[x^2-\frac{m-5}{m-1}x-\frac{2(m-3)}{(m-1)^2}]^{-3/2}(2x-\frac{m-5}{m-1})^2; \\
&+ [x^2-\frac{m-5}{m-1}x-\frac{2(m-3)}{(m-1)^2}]^{-1/2} \\
&= -\frac{1}{4}[x^2-\frac{m-5}{m-1}x-\frac{2(m-3)}{(m-1)^2}]^{-3/2}
\end{aligned}$$

$$h'''(x) = \frac{3}{8}[x^2-\frac{m-5}{m-1}x-\frac{2(m-3)}{(m-1)^2}]^{-5/2}(2x-\frac{m-5}{m-1}).$$

By the Taylor expansion,

$$\begin{aligned}
h(S) &= h(ES)+(S-ES)h'(ES)+\frac{1}{2}(S-ES)^2h''(ES) \\
&+ \frac{1}{2}(S-ES)^3\int_0^1(1-\lambda)^2h'''[\lambda S+(1-\lambda)ES]d\lambda.
\end{aligned}$$

Noting that $(S-\frac{m-3}{m-1})(S+\frac{2}{m-1}) \geq \frac{2}{m-1}S$, for $S > 1$ and $P(S < 1) = O(m^{-r})$ for any arbitrary $r > 0$ from (3.23), it follows now from (3.46) that

$$h'''[\lambda S+(1-\lambda)ES] \leq |h'''[\lambda S+(1-\lambda)ES]|I_{[S>1]}+O(m^{-r})$$

$$
\begin{aligned}
&\le \frac{3}{8}\{\frac{2}{m-1}(\lambda S+(1-\lambda)ES\}^{-5/2}2\{\lambda S+(1-\lambda)ES\}I_{[S>1]}\\
&\le \frac{3}{2(m-1)}[\lambda S+(1-\lambda)ES]^{-3/2}I_{[S>1]}\\
&\le \frac{3}{2(m-1)}[(1-\lambda)ES]^{-3/2}.
\end{aligned}
$$

Since $\int_0^1(1-\lambda)^2(1-\lambda)^{-3/2}d\lambda<\infty$ and $E|S-ES|^3=O(m^{-3/2})$, choosing $r>3/2$, it follows that

$$
E[(S-ES)^3\int_0^1(1-\lambda)^2h'''[\lambda S+(1-\lambda)ES]d\lambda]=O(m^{-3/2}).
$$

Also since $E(S)=B^{-1}$, it follows from

$$
\begin{aligned}
E[h(S)] &= (B^{-1}-\frac{m-3}{m-1})^{1/2}(B^{-1}+\frac{2}{m-1})^{1/2}\\
&- \frac{B^{-2}}{4(m-1)}(B^{-1}-\frac{m-3}{m-1})^{-3/2}(B^{-1}+\frac{2}{m-1})^{-3/2}+O(m^{-3/2})\\
&= (\frac{1-B}{B}+\frac{2}{m-1})^{1/2}(\frac{1}{B}+\frac{2}{m-1})^{1/2}\\
&- \frac{1}{4(m-1)B^2}(\frac{1-B}{B}+\frac{2}{m-1})^{-3/2}(\frac{1}{B}+\frac{2}{m-1})^{-3/2}+O(m^{-3/2})\\
&= \frac{(1-B)^{1/2}}{B}[1+\frac{B}{(m-1)(1-B)}+\frac{B}{m-1}]\\
&- \frac{B^3(1-B)^{-3/2}}{4(m-1)B^2}+O(m^{-3/2})\\
&= (1-B)^{1/2}B^{-1}+\frac{1}{m}[(2-B)(1-B)^{-1/2}-\frac{1}{4}(1-B)^{-3/2}B]\\
&+ O(m^{-3/2}). \qquad (3.47)
\end{aligned}
$$

Now combining (3.45) and (3.47), one gets the result.

CHAPTER 4
RANDOM EFFECTS NORMAL ANOVA MODEL WITH BALANCED LOSS FUNCTION

4.1 Introduction

In Chapter 3 of this dissertation, we developed the general algorithm for finding constrained Bayes and empirical Bayes estimators under balanced loss functions. In particular, for the normal-normal example, we found the Bayes risks of the empirical Bayes estimators correct up to order $O(m^{-1})$ (m being the number of cells in ANOVA problem) assuming the sample variances to be known, but the prior means and variances to be unknown.

In the present chapter, we derive constrained empirical Bayes estimators in the normal-normal set up when the sample variances are also unknown. First in Section 2, we find the constrained Bayes estimators and the Bayes risks of these estimators correct up to order $O(m^{-1})$ under the present set up. The results are slight extensions of those in Section 3.2. Next in Section 3, we find the constrained empirical Bayes estimators and the Bayes risks of these estimators also correct up to $O(m^{-1})$ under the present set up.

4.2 Constrained Bayes Estimators in the Balanced ANOVA Model

In this section we develop the constrained Bayes estimators in the balanced normal ANOVA model under balanced loss function and find also the Bayes risk of the such estimators valid up to $O(m^{-1})$.

4.2.1 Constrained Bayes Estimators

Consider the balanced normal ANOVA model with $Y_{ij} = \theta_i + e_{ij}$ and $\theta_i = \mu + \alpha_i$ ($j = 1, ..., k$; $i = 1, ...m$). Here the α_i and the e_{ij} are mutually independent with

$\alpha_i \stackrel{iid}{\sim} N(0, \tau^2)$ and $e_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$. Alternatively, in a Bayesian framework, this amounts to saying that $Y_{ij}|\theta_i \stackrel{iid}{\sim} N(\theta_i, \sigma^2)$, $i = 1, ...m$ and $\theta_i \stackrel{iid}{\sim} N(\mu, \tau^2)$.

Minimal sufficiency consideration allows us to restrict to $(X_1, ..., X_m, SSW)$, where $X_i = \frac{1}{k}\sum_{j=1}^{k} Y_{ij} = \bar{Y}_i$ and $SSW = \sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij} - \bar{Y}_i)^2$. We may note that marginally $X_1, ..., X_m$ and SSW are mutually independent with $X_i \stackrel{iid}{\sim} N(\mu, \tau^2 + \sigma^2/k)$, i.e., $N(\mu, \sigma^2/(kB))$, where $B = \frac{\sigma^2/k}{\sigma^2/k+\tau^2} = \frac{\sigma^2}{\sigma^2+k\tau^2}$ and $SSW \sim \sigma^2\chi^2_{m(k-1)}$.

From the results of the previous chapter, the constrained Bayes estimator of $\boldsymbol{\theta} = (\theta_1, ..., \theta_m)^T$ is given by

$$\hat{\boldsymbol{\theta}}^{CB} = a(\boldsymbol{X})(1 - B)(\boldsymbol{X} - \bar{X}\mathbf{1}_m) + \{(1 - B)\bar{X} + B\mu\}\mathbf{1}_m,$$

where $\boldsymbol{X} = (X_1, ..., X_m)^T$, $\bar{X} = m^{-1}\sum_{i=1}^{m} X_i$, and $\mathbf{1}_m$ is an m-component column vector with each element equal to 1. Also $a^2(X) = 1 + \frac{H_1(X)}{H_2(X)}$, where $H_1(X) = (m - 1)(1 - B)\sigma^2/k$ and $H_2(X) = (1 - B)^2\sum_{i=1}^{m}(X_i - \bar{X})^2 = (1 - B)^2(SSB/k)$, (say). Then, on simplification,

$$a^2(X) = 1 + \frac{\sigma^2}{(1 - B)MSB},$$

where $MSB = SSB/(m - 1)$.

4.2.2 Bayes Risk of Constrained Bayes Estimators

The calculation of the Bayes risk of the CB estimator is analogous to that in the previous chapter with minor modifications. For completeness, we outline the major steps. For the balanced loss as introduced in Chapter 3, the Bayes risk of $\hat{\boldsymbol{\theta}}^{CB}$ under the given model is

$$r(\hat{\boldsymbol{\theta}}^{CB}) = m^{-1}\{wE||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CB}||^2 + (1 - w)E||\hat{\boldsymbol{\theta}}^{CB} - \boldsymbol{\theta}||^2\}, \tag{4.1}$$

where $\boldsymbol{X}^T = (X_1, X_2, ..., X_m)$. We now find the Bayes risk of $\hat{\boldsymbol{\theta}}^{CB}$ in the following theorem.

<u>Theorem 4.1</u> Under the loss given in (4.1), the Bayes risk of $\hat{\boldsymbol{\theta}}^{CB}$ is given by

$$(1-w)(1-B)+\frac{\sigma^2}{k}a_1(B)+\frac{\sigma^2}{km}[a_2(B)-a_1(B)]+o(m^{-1}),$$

where $a_1(B)=B^{-1}\{1-(1-w)B\}[2-B-2(1-B)^{1/2}]$ and $a_2(B)=\frac{B}{2}(1-B)^{1/2}(1-(1-w)B)$.

<u>Proof of Theorem 4.1</u> As before, let $\boldsymbol{e}^{PM}=(1-B)\boldsymbol{X}+B\mu\mathbf{1}_m$. Then

$$\begin{aligned} m^{-1}E\|\hat{\boldsymbol{\theta}}^{CB}-\boldsymbol{\theta}\|^2 &= m^{-1}E\|\hat{\boldsymbol{\theta}}^{CB}-\boldsymbol{e}^{PM}+\boldsymbol{e}^{PM}-\boldsymbol{\theta}\|^2 \\ &= 1-B+m^{-1}E\|\boldsymbol{e}^{PM}-\hat{\boldsymbol{\theta}}^{CB}\|^2. \end{aligned} \tag{4.2}$$

But

$$\begin{aligned} \boldsymbol{e}^{PM}-\hat{\boldsymbol{\theta}}^{CB} &= (1-B)(\boldsymbol{X}-\bar{X}\mathbf{1}_m)-a(\boldsymbol{X})(1-B)(\boldsymbol{X}-\bar{X}\mathbf{1}_m) \\ &= (1-a(\boldsymbol{X}))(1-B)(\boldsymbol{X}-\bar{X}\mathbf{1}_m). \end{aligned} \tag{4.3}$$

So by (4.3), one gets

$$m^{-1}E\|\boldsymbol{e}^{PM}-\hat{\boldsymbol{\theta}}^{CB}\|^2=(1-B)^2\frac{m-1}{km}E[(1-a(\boldsymbol{X}))^2MSB]. \tag{4.4}$$

Next

$$\boldsymbol{X}-\hat{\boldsymbol{\theta}}^{CB}=(1-a(\boldsymbol{X})(1-B))(\boldsymbol{X}-\bar{X}\mathbf{1}_m)+B(\bar{X}-\mu)\mathbf{1}_m. \tag{4.5}$$

Hence, by (4.5), one gets

$$\begin{aligned} m^{-1}E\|\boldsymbol{X}-\hat{\boldsymbol{\theta}}^{CB}\|^2 &= m^{-1}E[(1-a(\boldsymbol{X})(1-B))^2\frac{m-1}{k}MSB] \\ &+ B^2(km)^{-1}(\sigma^2+k\tau^2) \\ &= \frac{m-1}{km}E[(1-a(\boldsymbol{X})(1-B))^2MSB]+\frac{B\sigma^2}{km}. \end{aligned} \tag{4.6}$$

Combining the results from (4.2), (4.4) and (4.6),

$$\begin{aligned} E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{CB})] &= (1-w)(1-B) + \frac{wB\sigma^2}{km} \\ &+ \frac{m-1}{km} E[(w(1-a(\boldsymbol{X})(1-B))^2 MSB] \\ &+ \frac{m-1}{km} E[(1-w)(1-B)^2(1-a(\boldsymbol{X}))^2 MSB]. \end{aligned} \tag{4.7}$$

On simplification,

$$\begin{aligned} & w\{1-(1-B)a(\boldsymbol{X})\}^2 + (1-w)(1-B)^2(1-a(\boldsymbol{X}))^2 \\ & = (1-B)^2 a^2(\boldsymbol{X}) - 2(1-B)(1-(1-w)B)a(\boldsymbol{X}) \\ & +\{w + (1-w)(1-B)^2\}. \end{aligned} \tag{4.8}$$

We now calculate

$$\begin{aligned} E[a^2(\boldsymbol{X})MSB] &= E[(1 + \frac{\sigma^2}{(1-B)MSB})MSB)] \\ &= E[MSB + \frac{\sigma^2}{1-B}] \\ &= (\sigma^2 + k\tau^2) + \frac{\sigma^2}{1-B} \\ &= \frac{\sigma^2}{B} + \frac{\sigma^2}{1-B} \\ &= \sigma^2 B^{-1}(1-B)^{-1}. \end{aligned} \tag{4.9}$$

Hence, from (4.8) and (4.9),

$$\begin{aligned} & E[\{w\{1-(1-B)a(\boldsymbol{X})\}^2 + (1-w)(1-B)^2(1-a(\boldsymbol{X}))^2\}MSB] \\ & = B^{-1}(1-B)\sigma^2 + \{w + (1-w)(1-B)^2\}B^{-1}\sigma^2 \\ & -2(1-B)(1-(1-w)B)E[a(\boldsymbol{X})MSB] \\ & = B^{-1}(1-B)\sigma^2 + B^{-1}\sigma^2 - 2(1-w)\sigma^2 + (1-w)B\sigma^2 \\ & -2(1-B)(1-(1-w)B)E[a(\boldsymbol{X})MSB] \\ & = \sigma^2(2-B)B^{-1}[1-(1-w)B] \end{aligned}$$

$$-2(1-B)(1-(1-w)B)E[a(\boldsymbol{X})MSB]. \tag{4.10}$$

Next we find

$$\begin{aligned}
E[a(\boldsymbol{X})MSB] &= E[(1+\frac{\sigma^2}{(1-B)MSB})^{1/2}MSB)] \\
&= (1-B)^{-1/2}E[(1-B)(MSB)^2+\sigma^2 MSB]^{1/2} \\
&= (1-B)^{-1/2}E[g(MSB)] \quad (say), \qquad (4.11)
\end{aligned}$$

where $g(x) = [(1-B)x^2+\sigma^2 x]^{1/2}$. By Taylor expansion, noting that $E(MSB) = \sigma^2/B$,

$$\begin{aligned}
g(MSB) &= g(\sigma^2/B)+(MSB-\sigma^2/B)g'(\sigma^2/B)+\frac{1}{2}(MSB-\sigma^2/B)^2 g''(\sigma^2/B) \\
&+ \frac{1}{2}(MSB-\sigma^2/B)^3\int_0^1(1-\lambda)^2 g'''[\lambda(MSB)+(1-\lambda)(\sigma^2/B)]d\lambda (4.12)
\end{aligned}$$

Now

$$g'(x) = \frac{1}{2}\{(1-B)x^2+\sigma^2 x\}^{-1/2}\{2(1-B)x+\sigma^2\};$$

$$\begin{aligned}
g''(x) &= -\frac{1}{4}\{(1-B)x^2+\sigma^2 x\}^{-3/2}\{2(1-B)x+\sigma^2\}^2 \\
&+ (1-B)\{(1-B)x^2+\sigma^2 x\}^{-1/2} \\
&= -\frac{\sigma^4}{4}\{(1-B)x^2+\sigma^2 x\}^{-3/2};
\end{aligned}$$

$$g'''(x) = \frac{3\sigma^4}{8}\{(1-B)x^2+\sigma^2 x\}^{-5/2}\{2(1-B)x+\sigma^2\}.$$

Hence,

$$\begin{aligned}
g(\sigma^2/B) &= [(1-B)(B^{-1}\sigma^2)^2+B^{-1}\sigma^4]^{1/2} \\
&= [B^{-1}\sigma^4\{(1-B)B^{-1}+1\}]^{1/2} \\
&= B^{-1}\sigma^2; \qquad (4.13)
\end{aligned}$$

$$g''(\sigma^2/B) = -\frac{\sigma^4}{4}(B^{-2}\sigma^4)^{-3/2} = -\frac{B^3}{4\sigma^2}; \tag{4.14}$$

Finally, since $x > 0$,

$$\begin{aligned}
|g'''(x)| &\leq \frac{3\sigma^4(\sigma^2+2x)}{8x^{5/2}\sigma^5(1-B)^{5/2}} \\
&\leq \frac{3\sigma^4(\sigma^2+2x)}{8x^{5/2}\sigma^5} \\
&= \frac{3(\sigma^2+2x)}{8\sigma x^{5/2}}
\end{aligned}$$

Thus,

$$\begin{aligned}
|g'''[\lambda X + (1-\lambda)EX]| &\leq \frac{3\sigma}{8}[\lambda X + (1-\lambda)EX]^{-5/2} \\
&\quad + \frac{3}{4\sigma}[\lambda X + (1-\lambda)EX]^{-3/2} \\
&\leq \frac{3\sigma}{8}(1-\lambda)^{-5/2}B^{5/2}\sigma^{-5} \\
&\quad + \frac{3}{4\sigma}(1-\lambda)^{-3/2}B^{3/2}\sigma^{-3}.
\end{aligned}$$

Hence, $\int_0^1(1-\lambda)^2|g'''[\lambda X + (1-\lambda)EX]|d\lambda < \infty$. Also we know that $E|MSB - \sigma^2/B|^3 = O(m^{-3/2})$ and

$$E(MSB - \sigma^2/B)^2 = \frac{2B^{-2}\sigma^4}{m-1}. \tag{4.15}$$

Combining the results from (4.12) to (4.15),

$$\begin{aligned}
E[g(MSB)] &= B^{-1}\sigma^2 - \frac{B^{-2}\sigma^4}{m-1}\left(\frac{B^3}{4\sigma^2}\right) + O(m^{-3/2}) \\
&= B^{-1}\sigma^2 - \frac{B\sigma^2}{4m} + O(m^{-3/2}).
\end{aligned} \tag{4.16}$$

Thus, by (4.10), (4.11) and (4.16),

$$\begin{aligned}
E[L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}^{CB})] &= (1-w)(1-B) + \frac{wB\sigma^2}{km} + \\
&+ \frac{(m-1)\sigma^2}{km}(2-B)B^{-1}\{1-(1-w)B\} \\
&- \frac{2(m-1)\sigma^2}{km}(1-B)^{1/2}(1-(1-w)B)\left(B^{-1} - \frac{B}{4m} + O(m^{-3/2})\right)
\end{aligned}$$

$$
\begin{aligned}
&= (1-w)(1-B) \\
&+ \frac{\sigma^2}{k}B^{-1}\{1-(1-w)B\}[2-B-2(1-B)^{1/2}] \\
&- \frac{\sigma^2}{km}[B^{-1}\{1-(1-w)B\}[2-B-2(1-B)^{1/2}] \\
&+ \frac{\sigma^2}{km}[\frac{1}{2}B(1-B)^{1/2}\{1-(1-w)B\}]+o(m^{-1}) \\
&= (1-w)(1-B)+\frac{\sigma^2}{k}a_1(B)+\frac{\sigma^2}{km}[a_2(B)-a_1(B)] \\
&+ o(m^{-1}),
\end{aligned}
$$

where $a_1(B) = B^{-1}\{1-(1-w)B\}[2-B-2(1-B)^{1/2}]$ and $a_2(B) = \frac{B}{2}(1-B)^{1/2}(1-(1-w)B)$.

The following section develops comstrained EB estimators and finds their Bayes risks order up to $O(m^{-1})$.

4.3 Constrained Empirical Bayes Estimators in the Balanced ANOVA Model

The constrained empirical Bayes (EB) estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}^{CEB} = a_{EB}(\boldsymbol{X})(1-\hat{B})(\boldsymbol{X}-\bar{X}\mathbf{1}_m)+\bar{X}\mathbf{1}_m,$$

after substitution of μ by $\bar{X}$, B by $\hat{B}$ and σ^2 by MSW. Here $\hat{B} = min\{\frac{m-3}{m-1}, \frac{(m-3)MSW}{(m-1)MSB}\}$ and $a_{EB}(\boldsymbol{X}) = 1+\frac{MSW}{(1-\hat{B})MSB}$. We now find the Bayes risk of $\hat{\boldsymbol{\theta}}^{CEB}$ in the following theorem.

<u>Thorem 4.2</u> Under the loss given in (4.1), the Bayes risk of $\hat{\boldsymbol{\theta}}^{CEB}$ in the balanced ANOVA model is given by

$$(1-w)(1-B)+\frac{\sigma^2}{k}a_1(B)+\frac{\sigma^2}{km}[(1-w)B-a_1(B)+a_2(B)+a_3(B)]+o(m^{-1}),$$

where $a_1(B) = (1-(1-w)B)B^{-1}(2-B+2(1-B)^{1/2}]$, $a_2(B) = 2(2-B)(1-2(1-B)^{1/2}$ and $a_3(B) = \frac{B}{k-1}(\frac{4}{k}-\frac{k(1-B)^{-3/2}}{2})$.

<u>Proof of Theorem 4.2</u> We begin calculating

$$||\boldsymbol{X}-\hat{\boldsymbol{\theta}}^{CEB}||^2 = (1-a_{EB}(\boldsymbol{X})(1-\hat{B}))(\boldsymbol{X}-\bar{X}\mathbf{1}_m)||^2$$

$$\begin{aligned}
&= [(1 - a_{EB}(\boldsymbol{X})(1-\hat{B}))^2 \frac{m-1}{k} MSB] \\
&= \frac{m-1}{k}[MSB\{1 + (1-\hat{B})^2 + (1-\hat{B})\frac{MSW}{MSB}\}] \\
&- \frac{2(m-1)}{k} E[a_{EB}(\boldsymbol{X})(1-\hat{B})MSB] \\
&= \frac{m-1}{k}[MSB + (1-\hat{B})^2 MSB + (1-\hat{B})MSW] \\
&- \frac{2(m-1)}{k} E[a_{EB}(\boldsymbol{X})(1-\hat{B})MSB].
\end{aligned}$$

Hence,

$$\begin{aligned}
m^{-1}E\|\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}\|^2 &= \frac{m-1}{km}B^{-1}\sigma^2 + \frac{m-1}{km}E[(1-\hat{B})^2 MSB] \\
&+ \frac{m-1}{km}E[(1-\hat{B})MSW] \\
&- \frac{2(m-1)}{km}E[a_{EB}(x)(1-\hat{B})MSB]. \qquad (4.17)
\end{aligned}$$

Define $\hat{\hat{B}} = \frac{(m-3)MSW}{(m-1)MSB}$. First we calculate $E[(1-\hat{\hat{B}})^2 MSB]$. Noting that MSB and MSW are independently distributed with $MSB \sim (\sigma^2/B)\chi^2_{m-1}/(m-1)$ and $MSW \sim \sigma^2\chi^2_{m(k-1)}/m(k-1)$, one gets

$$\begin{aligned}
E[(1-\hat{\hat{B}})^2 MSB] &= E[MSB - 2\frac{m-3}{m-1}MSW + \frac{(m-3)^2}{(m-1)^2}\frac{(MSW)^2}{MSB}] \\
&= \sigma^2/B - 2\frac{m-3}{m-1}\sigma^2 + \frac{(m-3)^2}{(m-1)^2}\frac{m(k-1)+2}{m(k-1)}\frac{m-1}{m-3}B \\
&= \sigma^2[\frac{1}{B} - 2(1-\frac{2}{m-1}) + (1-\frac{2}{m-1})(1+\frac{2}{m(k-1)})B] \\
&= \sigma^2[\frac{(1-B)^2}{B} + \frac{4}{m-1} - (\frac{2}{m-1} - \frac{2}{m(k-1)})B \\
&\quad +O(m^{-2})]. \qquad (4.18)
\end{aligned}$$

Next we show that

Lemma 4.1 For any arbitrary $r > 0$,

$$E[(\hat{\hat{B}} - \hat{B})^2 MSB] = o(m^{-r}).$$

Proof of Lemma 4.1

$$\begin{aligned} E[(\hat{\hat{B}} - \hat{B})^2 MSB] &= E[(\frac{m-3}{m-1})^2 (1 - \frac{MSW}{MSB})^2 I_{MSW>MSB}] \\ &\le E^{1/2}(1 - \frac{MSW}{MSB})^4 P^{1/2}(MSW > MSB) \\ &= E^{1/2}(1 - \frac{MSW}{MSB})^4 P^{1/2}(F_{m(k-1),m-1} > B^{-1}). \quad (A) \end{aligned}$$

But

$$\begin{aligned} P^{1/2}(F_{m(k-1),m-1} > B^{-1}) &= P^{1/2}(F_{m(k-1),m-1} - \frac{m-1}{m-3} > B^{-1} - \frac{m-1}{m-3}) \\ &\le \frac{E[F_{m(k-1),m-1} - (m-1)/(m-3)]^{2r}}{(B^{-1} - 1 - 2/(m-3))^{2r}}. \end{aligned}$$

We next show that

$$E[F_{m(k-1),m-1} - (m-1)/(m-3)]^{2r} = O(m^{-r}) \quad as \quad m \to \infty. \quad (B)$$

In order to prove (B), we begin with the result

$$\sqrt{m} \begin{bmatrix} MSB - \sigma^2/B \\ MSW - \sigma^2 \end{bmatrix} \to^d N[0, Diag(2\sigma^4/B^2, 2\sigma^4)],$$

which follows as a consequence of the central limit theorem and the independence of MSB and MSW. Hence, by the Delta method,

$$\sqrt{m}(\frac{MSW}{MSB} - B) \to^d N(0, B^2).$$

This shows that $\sqrt{m}[F_{m(k-1),m-1} - 1] = O_p(1)$, since $MSW/MSB \sim BF_{m(k-1),m-1}$. Further, for any $r \ge 1$,

$$sup_{m \ge 16r+1} m^{2r} E[F_{m(k-1),m-1} - 1]^{4r} = sup_{m \ge 16r+1} m^{2r} E\left[\frac{U_{1m} - U_{2m}}{U_{2m}}\right]^{4r},$$

where U_{1m} and U_{2m} are independent with $U_{1m} \sim \chi^2_{m(k-1)}/m(k-1)$ and $U_{2m} \sim \chi^2_{m-1}/(m-1)$. By the Schwarz inequality,

$$
\begin{aligned}
E[U_{2m}^{-1}(U_{1m}-U_{2m})]^{4r} &\leq E^{1/2}(U_{2m}^{-8r})E^{1/2}[(U_{1m}-1)-(U_{2m}-1)]^{8r} \\
&= O(1)O(m^{-2r}),
\end{aligned}
$$

since by c_δ-inequality, for $r \geq 1$,

$$
E[(U_{1m}-1)-(U_{2m}-1)]^{8r} \leq 2^{8r-1}E[(U_{1m}-1)^{8r}-(U_{2m}-1)^{8r}] = O(m^{-4r}),
$$

and for $m \geq 16r+1$, $E(U_{2m}^{-8r}) = O(1)$. Thus, $m^{2r}[F_{m(k-1),m-1}-1]^{4r}$ is uniformly integrable in $m \geq 16r+1$. Hence, the left hand side of (B) is $O(m^{-r})$ for $r > 0$. This implies that the left hand side of (A) equals to $O(1)O(m^{-r}) = O(m^{-r})$.

In view of the above lemma, it follows that

$$
\begin{aligned}
E[|\hat{\hat{B}}-\hat{B}|(1-\hat{B})MSB] &\leq E[|\hat{\hat{B}}-\hat{B}|MSB] \\
&\leq E^{1/2}[(\hat{\hat{B}}-\hat{B})^2E^{1/2}(MSB) \\
&= o(m^{-r}),
\end{aligned}
\tag{4.19}
$$

for an arbitrary $r > 0$. Thus, from (4.18) and (4.19),

$$
\begin{aligned}
E[(1-\hat{B})^2MSB] &= \sigma^2[\frac{(1-B)^2}{B}+\frac{4}{m-1}-(\frac{2}{m-1}-\frac{2}{m(k-1)})B] \\
&+ o(m^{-1}).
\end{aligned}
\tag{4.20}
$$

Next we find

$$
\begin{aligned}
E[(1-\hat{B})MSW] &= E[(1-\hat{\hat{B}})MSW+(\hat{\hat{B}}-\hat{B})MSW] \\
&= E[(1-\hat{\hat{B}})MSW]+o(m^{-r}),
\end{aligned}
\tag{4.21}
$$

for an arbitrary $r > 0$. Also

$$
\begin{aligned}
E[(1-\hat{\hat{B}})MSW] &= E[(1-\frac{m-3}{m-1}\frac{MSW}{MSB})MSW] \\
&= \sigma^2-\frac{m-3}{m-1}E[\frac{(MSW)^2}{MSB}]
\end{aligned}
$$

$$\begin{aligned}
&= \sigma^2 - \sigma^2 \frac{m-3}{m-1}\frac{m(k-1)+2}{m(k-1)}\frac{m-1}{m-3}B \\
&= \sigma^2[1 - B(1 - \frac{2}{m(k-1)}] \\
&= \sigma^2[1 - B + \frac{2B}{m(k-1)}]. \qquad (4.22)
\end{aligned}$$

Combining the results from (4.17), (4.20), (4.21) and (4.22), we get

$$\begin{aligned}
m^{-1}E||\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}||^2 &= \frac{(m-1)}{km}B^{-1}\sigma^2 \\
&+ \frac{(m-1)}{km}\sigma^2[\frac{(1-B)^2}{B} + \frac{4}{m-1} - (\frac{2}{m-1} - \frac{2}{m(k-1)})B] \\
&+ \frac{(m-1)}{km}\sigma^2[1 - B + \frac{2B}{m(k-1)}] \\
&- \frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})MSB] + o(m^{-1}) \\
&= \frac{\sigma^2}{k}[B^{-1} + \frac{(1-B)^2}{B} + 1 - B] \\
&+ \frac{\sigma^2}{km}[-\{B^{-1} + \frac{(1-B)^2}{B} + 1 - B\} + 4 - 2B + \frac{4B}{k(k-1)}] \\
&- \frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})MSB] + o(m^{-1}) \\
&= \frac{\sigma^2}{k}(2-B)B^{-1} + \frac{\sigma^2}{km}[(2-B)(2-B^{-1}) + \frac{4B}{k(k-1)}] \\
&- \frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})MSB] + o(m^{-1}). \qquad (4.23)
\end{aligned}$$

Next we find

$$\begin{aligned}
m^{-1}E||\boldsymbol{\theta}^{CEB} - \boldsymbol{\theta}||^2 &= m^{-1}E||\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{e}^{PM} + \boldsymbol{e}^{PM} - \boldsymbol{\theta}||^2 \\
&= (1-B) + m^{-1}E||\boldsymbol{e}^{PM} - \hat{\boldsymbol{\theta}}^{CEB}||^2. \qquad (4.24)
\end{aligned}$$

But

$$\begin{aligned}
\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{e}^{PM} &= a_{EB}(\boldsymbol{X})(1-\hat{B})(\boldsymbol{X} - \bar{X}\mathbf{1}_m) - (1-B)\boldsymbol{X} - B\mu\mathbf{1}_m \\
&= [a_{EB}(\boldsymbol{X})(1-\hat{B}) - (1-B)](\boldsymbol{X} - \bar{X}\mathbf{1}_m) \\
&+ B(\bar{X} - \mu)\mathbf{1}_m.
\end{aligned}$$

So by the independence of $\bar{X}$ and $\boldsymbol{X} - \bar{X}\mathbf{1}_m$,

$$\begin{aligned} m^{-1}E\|\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{e}^{PM}\|^2 &= \frac{B\sigma^2}{km} \\ &+ \frac{m-1}{km}E[\{a_{EB}(\boldsymbol{X})(1-\hat{B}) - (1-B)\}^2 MSB]. \end{aligned}$$

Now

$$\begin{aligned} &\frac{m-1}{km}E[\{a_{EB}(\boldsymbol{X})(1-\hat{B} - (1-B)\}^2 MSB] \\ &= \frac{m-1}{km}E[\{a_{EB}^2(\boldsymbol{X})(1-\hat{B})^2 + (1-B)^2\}MSB] \\ &-\frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})(1-B)MSB] \\ &= \frac{m-1}{km}E[(1-\hat{B})^2 MSB + (1-\hat{B})MSW + (1-B)^2 MSB] \\ &-\frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})(1-B)MSB] \\ &= \frac{(m-1)}{km}\sigma^2[\frac{(1-B)^2}{B} + \frac{4}{m-1} - (\frac{2}{m-1} - \frac{2}{m(k-1)})B] \\ &+\frac{(m-1)}{km}\sigma^2[1-B+\frac{2B}{m(k-1)}] + \frac{(m-1)}{km}\sigma^2\frac{(1-B)^2}{B} \\ &-\frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})(1-B)MSB] \\ &= \frac{\sigma^2}{k}(1-B)(2-B)B^{-1} \\ &+\frac{\sigma^2}{km}[(2-B)(3-B^{-1}) + \frac{4B}{k(k-1)}] \\ &-\frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})(1-B)MSB] + o(m^{-1}). \end{aligned}$$

Hence, (4.24) can be expressed as

$$\begin{aligned} m^{-1}E\|\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{\theta}\|^2 &= (1-B) + \frac{B\sigma^2}{km} + \frac{\sigma^2}{k}(1-B)(2-B) \\ &+ \frac{\sigma^2}{km}[(1+B)(2-B) + \frac{4B}{k(k-1)}] \\ &- \frac{2(m-1)}{km}E[a_{EB}(\boldsymbol{X})(1-\hat{B})(1-B)MSB] + o(m^{-1}) \\ &= (1-B) + \frac{\sigma^2}{k}(1-B)(2-B) \\ &+ \frac{\sigma^2}{km}[B + (1+B)(2-B) + \frac{4B}{k(k-1)}] \end{aligned}$$

$$\begin{aligned} &- \frac{2(m-1)}{km} E[a_{EB}(\boldsymbol{X})(1-\hat{B})(1-B)MSB] \\ &+ o(m^{-1}). \end{aligned} \tag{4.25}$$

Combining (4.23) and (4.25) with some simplifications, we get

$$\begin{aligned} &m^{-1}\{wE\|\boldsymbol{X} - \hat{\boldsymbol{\theta}}^{CEB}\|^2 + (1-w)E\|\hat{\boldsymbol{\theta}}^{CEB} - \boldsymbol{\theta}\|^2\} \\ &= (1-w)(1-B) + \frac{\sigma^2}{k}[(1-(1-w)B)B^{-1}(2-B)] \\ &+ \frac{\sigma^2}{km}[(2-B)(3-B^{-1}-w) + (1-w)B + \frac{4B}{k(k-1)}] \\ &- \frac{2(m-1)}{km}(1-(1-w)B)E[(1-\hat{B})a_{EB}MSB] \\ &+ o(m^{-1}). \end{aligned} \tag{4.26}$$

Next we find

$$\begin{aligned} E[a_{EB}(1-\hat{B})MSB] &= E[(1-\hat{B})^2 MSB^2 + (1-\hat{B})(MSB)(MSW)]^{1/2} \\ &= E[\{(1-\hat{\hat{B}})^2(MSB)^2 + (1-\hat{\hat{B}})(MSB)(MSW)\}^{1/2} I_{[\hat{B}=\hat{\hat{B}}]}] \\ &+ E[\{(1-\frac{m-3}{m-1})^2(MSB)^2 + (1-\frac{m-3}{m-1})(MSB)(MSW)\}^{1/2} \\ &\times I_{[\hat{B}\neq\hat{\hat{B}}]}]. \end{aligned} \tag{4.27}$$

On simplification,

$$\begin{aligned} &(1-\hat{\hat{B}})^2(MSB)^2 + (1-\hat{\hat{B}})(MSB)(MSW) \\ =\ & [1-\frac{(m-3)MSW}{(m-1)MSB}]^2(MSB)^2 + [1-\frac{(m-3)MSW}{(m-1)MSB}](MSB)(MSW) \\ =\ & (MSB)^2 - \frac{(m-3)(MSW)(MSB)}{(m-1)} + \frac{(m-3)^2(MSW)^2}{(m-1)^2} \\ +\ & (MSB)(MSW) - \frac{m-3}{m-1}(MSW)^2 \\ =\ & (MSB)^2 - \frac{m-5}{m-1}(MSB)(MSW) - \frac{2(m-3)}{(m-1)^2}(MSW)^2 \\ =\ & g^2(MSB, MSW) \quad (say), \end{aligned}$$

where $g(x,y) = [x^2 - \frac{m-5}{m-1}xy - \frac{2(m-3)}{(m-1)^2}y^2]^{1/2} = (x - \frac{m-3}{m-1}y)^{1/2}(x + \frac{2}{m-1}y)^{1/2}$, $y \leq x$.

By Taylor expansion, for $y \le x$,

$$\begin{aligned} g(x,y) &= g(\sigma^2/B,\sigma^2) \\ &+ (x-\sigma^2/B)\frac{\partial g}{\partial x}|_{x=\sigma^2/B,y=\sigma^2} + (y-\sigma^2)\frac{\partial g}{\partial y}|_{x=\sigma^2/B,y=\sigma^2} \\ &+ \frac{1}{2}[(x-\sigma^2/B)^2\frac{\partial^2 g}{\partial x^2}|_{x=\sigma^2/B,y=\sigma^2} + (y-\sigma^2)^2\frac{\partial^2 g}{\partial y^2}|_{x=\sigma^2/B,y=\sigma^2}] \\ &+ (x-\sigma^2/B)(y-\sigma^2)\frac{\partial^2 g}{\partial x \partial y}|_{x=\sigma^2/B,y=\sigma^2} + R(MSB,MSW). \end{aligned} \tag{4.28}$$

As in the previous chapters, we can show that

$$E[|R(MSB,MSW)|I_{[\hat{B}=\hat{\hat{B}}]}] \le E[|R(MSB,MSW)|] = O(m^{-3/2}). \tag{4.29}$$

Also, for $y < x$,

$$\begin{aligned} \frac{\partial g}{\partial x} &= \frac{1}{g(x,y)}(x - \frac{(m-5)y}{2(m-1)}); \\ \frac{\partial g}{\partial y} &= -\frac{1}{2(m-1)g(x,y)}[(m-5)x + \frac{4(m-3)y}{m-1}]; \\ \frac{\partial^2 g}{\partial x^2} &= \frac{1}{g(x,y)} - \frac{1}{g^2(x,y)}(x - \frac{(m-5)y}{2(m-1)})\frac{\partial g}{\partial x}; \\ &= \frac{1}{g(x,y)} - \frac{1}{g^3(x,y)}(x - \frac{(m-5)y}{2(m-1)})^2 \\ &= -\frac{1}{g^3(x,y)}[\frac{8(m-3)y^2 + (m-5)^2y^2}{4(m-1)^2})] \\ &= -\frac{y^2}{4g^3(x,y)}; \\ \frac{\partial^2 g}{\partial y^2} &= -\frac{1}{2(m-1)}\left[\frac{4(m-3)}{(m-1)g(x,y)} - \frac{1}{g^2(x,y)}((m-5)x + \frac{4(m-3)y}{m-1})\frac{\partial g}{\partial y}\right]; \\ &= -\frac{1}{2(m-1)}\left[\frac{4(m-3)}{(m-1)g(x,y)} - \frac{1}{g^3(x,y)}((m-5)x + \frac{4(m-3)y}{m-1})^2\right] \\ &= -\frac{1}{4(m-1)^2}\left[\frac{8(m-3) + (m-5)^2}{g^3(x,y)}x^2\right] \\ &= -\frac{x^2}{4g^3(x,y)}; \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 g}{\partial x \partial y} &= \frac{1}{g(x,y)}\left(-\frac{m-5}{2(m-1)}\right) - \frac{1}{g^2(x,y)}\left(x - \frac{(m-5)y}{2(m-1)}\right)\frac{\partial g}{\partial y} \\
&= -\frac{1}{2(m-1)}\left[\frac{m-5}{g(x,y)} + \left(x - \frac{(m-5)y}{2(m-1)}\right)\left((m-5)x + \frac{4(m-3)y)}{m-1}\right)/g^3(x,y)\right] \\
&= -\frac{1}{4(m-1)^2}\left[\frac{-2(m-5)^2 - 8(m-3) + (m-5)^2}{g^3(x,y)}xy\right] \\
&= \frac{xy}{4g^3(x,y)}. \qquad (4.30)
\end{aligned}$$

Hence, by (4.28), (4.29) and (4.30),

$$\begin{aligned}
E[g(x,y)] &= g(\sigma^2/B,\sigma^2) \\
&- \frac{1}{2}\left[E\left[\frac{(x-\sigma^2/B)^2\sigma^4}{4g^3(\sigma^2/B,\sigma^2)} + \frac{(y-\sigma^2)^2\sigma^4}{4g^3(\sigma^2/B,\sigma^2)B^2} - 2\frac{(x-\sigma^2/B)(y-\sigma^2)\sigma^4}{4g^3(\sigma^2/B,\sigma^2)B}\right]\right] \\
&+ O(m^{-3/2}) \\
&= g(\sigma^2/B,\sigma^2) - \frac{\sigma^4}{8g^3(\sigma^2/B,\sigma^2)}\left[E(x-\sigma^2/B)^2 + \frac{1}{B^2}E(y-\sigma^2)^2\right] \\
&+ O(m^{-3/2}) \\
&= g(\sigma^2/B,\sigma^2) - \frac{\sigma^4}{8g^3(\sigma^2/B,\sigma^2)}\left[\frac{2\sigma^4}{(m-1)B^2} + 2\sigma^4 m(k-1)B^2\right] \\
&+ O(m^{-3/2}) \\
&= g(\sigma^2/B,\sigma^2) - \left[\frac{\sigma^8}{4g^3(\sigma^2/B,\sigma^2)B^2}\right]\left[\frac{mk-1}{m(m-1)(k-1)}\right] \\
&+ O(m^{-3/2}) \qquad (4.31)
\end{aligned}$$

and so,

$$\begin{aligned}
g(\sigma^2/B,\sigma^2) &= \left(\frac{\sigma^2}{B} - \frac{(m-3)\sigma^2}{m-1}\right)^{1/2}\left(\frac{\sigma^2}{B} + \frac{2\sigma^2}{m-1}\right)^{1/2} \\
&= \sigma^2\left(\frac{1-B}{B} + \frac{2}{m-1}\right)^{1/2}\left(\frac{1}{B} + \frac{2}{m-1}\right)^{1/2} \\
&= \sigma^2(1-B)^{1/2}B^{-1}\left(1 + \frac{2B}{(m-1)(1-B)}\right)^{1/2}\left(1 + \frac{2B}{m-1}\right)^{1/2} \\
&= \sigma^2(1-B)^{1/2}B^{-1}\left[1 + \frac{2B(2-B)}{(m-1)(1-B)}\right] + o(m^{-1}). \qquad (4.32)
\end{aligned}$$

Combining (4.31) and (4.32),

$$\begin{aligned}
E[g(x,y)] &= (1-B)^{1/2}B^{-1}\sigma^2(1+\frac{2B(2-B)}{(m-1)(1-B)}) \\
&- \frac{\sigma^8(mk-1)}{4m(m-1)(k-1)B^2}\sigma^{-6}(1-B)^{-3/2}B^3 + o(m^{-1}) \\
&= (1-B)^{1/2}B^{-1}\sigma^2 \\
&+ \frac{\sigma^2}{m}[2(2-B)(1-B)^{1/2} - \frac{k(1-B)^{-3/2}B}{4(k-1)}] + o(m^{-1}) \\
&= h(\sigma^2/B,\sigma^2), \quad (say). \qquad (4.33)
\end{aligned}$$

Now

$$\begin{aligned}
E[g(MSB, MSW)I_{[\hat{B}=\hat{\hat{B}}]}] &= E[h(\sigma^2/B,\sigma^2)I_{[\hat{B}=\hat{\hat{B}}]}] + O(m^{-3/2}) \\
&= h(\sigma^2/B,\sigma^2)[1-P(\hat{B}\neq\hat{\hat{B}})] + O(m^{-3/2}) \\
&= h(\sigma^2/B,\sigma^2) + O(m^{-3/2}). \qquad (4.34)
\end{aligned}$$

Finally, it is easy to check that

$$\begin{aligned}
& E[\{(1-\frac{m-3}{m-1})^2(MSB)^2 + (1-\frac{m-3}{m-1})(MSB)(MSW)\}^{1/2}I_{[\hat{B}\neq\hat{\hat{B}}]}] \\
\leq\ & E^{1/2}[\{(1-\frac{m-3}{m-1})^2(MSB)^2 + (1-\frac{m-3}{m-1})(MSB)(MSW)\}^{1/2}]P^{1/2}(\hat{B}\neq\hat{\hat{B}}) \\
\leq\ & [E\{(MSB)^2 + (MSB)(MSW)\}]^{1/2}P^{1/2}(\hat{B}\neq\hat{\hat{B}}) \\
=\ & o(m^{-r}), \qquad (4.35)
\end{aligned}$$

for and arbitrary $r > 0$. Combining (4.26) with (4.33), (4.34) and (4.35), the proof is completed.

CHAPTER 5
CONCLUSION & FUTURE STUDY

5.1 Conclusion

One of the main objectives of this dissertation is to obtain the constrained Bayes and empirical Bayes estimators, and also as measures of precision, the asymptotic mean squared errors (MSE's) of these estimators which are correct up to a certain order. In addition, constrained Bayes and empirical Bayes estimators and their Bayes risks are found under both squared error loss and balanced loss functions. The asymptotic Bayes risks are calculated, and asymptotically unbiased estimators of these Bayes risks are obtained.

In Chapter 2, we considered the asymptotic expansion of the MSE of constrained James-Stein estimators. This expansion is valid up to $O(m^{-1})$, where m denotes the sample size. We also provided an estimator of the MSE asymptotically valid up to $O(m^{-1})$. A simulation study was undertaken to evaluate the performance of these estimators.

Chapter 3 developed constrained Bayes and empirical Bayes estimators under balanced loss functions. In particular, such estimators were derived under the one-parameter exponential family of distributions. In the normal-normal example, asymptotic expansions of MSE's of the Bayes and empirical Bayes estimators were provided which were asymptotically valid up to $O(m^{-1})$. In addition, similar asymptotic expansions of MSE's of constrained Bayes and empirical Bayes estimators were also provided. Estimators of these MSE's in the spirit of Chapter 2 were given.

Chapter 4 developed constrained Bayes and constrained empirical Bayes estimators for the random effects balanced normal ANOVA model when both

variance components were unknown. The asymptotic MSE's valid up to $O(m^{-1})$ were derived as in the previous chapters.

5.2 Future Study

This dissertation is devoted exclusively to balanced data, that is when the number of observations per cell is the same. We want to continue this research for unbalanced data, i.e., with varying number of observations in the different cells. We also want to extend our findings for cross-classificatory models with or without interaction, and develop similar constrained Bayes and empirical Bayes estimators both under squared error and balanced loss functions.

REFERENCES

Cressie, N. (1989). Empirical Bayes Estimation of Undercount in the Decennial Census. *Journal of the American Statistical Association*, 84, 1033-1044.

Datta G.S. and Lahiri, P. (2000). A Unified Measure of Uncertainty of Estimated Best Linear Unbiased Predictors in Small-area Estimation Problems. *Statistica Sinica*, 10, 613-628.

Efron, B. and Morris, C. (1973). Stein's Estimation Rule and Its Competitors–An Empirical Bayes Approach. *Journal of the American Statistical Association*, 68, 117-130.

Ghosh, M. and Meeden, G. (1986). Empirical Bayes Estimation in Finite Population Sampling. *Journal of the American Statistical Association*, 81, 1058-1062.

Ghosh, M. and Lahiri, P. (1987). Robust Empirical Bayes Estimation of Means from Stratified Samples. *Journal of the American Statistical Association*, 82, 1153-1162.

Ghosh, M. (1992). Constrained Bayes Estimation with Applications. *Journal of the American Statistical Association*, 87, 533-540.

Ghosh, M. and Kim, D. (2002). Multivariate Constrained Bayes Estimation. *Pakistan Journal of Statistics*, 18(2), 143-148.

James, W. and Stein, C. (1961). Estimation with Quadratic Loss. *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, 1, 361-380, Univ. California Press, Berkeley.

Lahiri, P. (1990). "Adjusted" Bayes and Empirical Bayes Estimation in Finite Population Sampling. *Sankhya, Indian Journal of Statistics*, Ser. B, 52, 50-66.

Lahiri, P. and Rao, J.N.K. (1995). Robust Estimation of Mean Squared Error of Small Area Estimators. *Journal of the American Statistical Association*, 90, 758-766.

Lindley, D.V. (1962). Discussion of Professor Stein's Paper. *Journal of the Royal Statistical Society*, Ser. B, 24, 265-296.

Louis, T.A. (1984). Estimating a Population of Parameter Values Using Bayes and Empirical Bayes Methods. *Journal of the American Statistical Association*, 79, 393-398.

Louis, T.A. (2001). Bayes/EB Ranking, Histogram and Parameter Estimation: Issues and Research Agenda. *Empirical Bayes and Likelihood Inference*, Springer-Verlag, New York, 1-16.

Morris, C. (1981). Parametric Empirical Bayes Confidence Intervals. In *Scientific Inference, Data Analysis, and Robustness*, eds. G.E.P. Box, T. Leonard and C.F. Jeff Wu, Academic Press, 25-50.

Morris, C. (1982). Natural Exponential Families with Quadratic Variance Functions. *Annals of Statistics*, 10, No. 1, 65-80

Morris, C. (1983). Natural Exponential Families with Quadratic Variance Functions: Statistical Theory. *Annals of Statistics*, 11, No. 2, 515-529.

Morris, C. (1983). Parametric Empirical Bayes Inference: Theory and Apllications. *Journal of the American Statistical Association*, 78, No. 381, Applications Section, 47-55.

Prasad, N.G.N. and Rao, J.N.K. (1990). The Estimation of the Mean Squared Error of Small-Area Estimators. *Journal of the American Statistical Association*, 85, 163-171.

Shen, W. and Louis, T.A. (1998). Triple-goal Estimates in Two-stage Hierarchical Models. *Journal of the Royal Statistical Society*, Ser. B, 60, 455-471.

Spjotvoll, E. and Thomsen, I. (1987). Application of Some Empirical Bayes Methods on Small Area Statistics. *Proceedings of the International Statistical Institute*, 2, 435-449.

Zellner A. (1988). Bayesian Analysis in Econometrics. *Journal of Econometrics*, 37, 27-50.

Zellner A. (1992). Bayesian and Non-Bayesian Estimation Using Balanced Loss Functions. *Statistical Decision Theory and Related Topics V*, Springer-Verlag, New York, 377-390.

BIOGRAPHICAL SKETCH

The author, Myung Joon Kim, was born on February 20, 1973, in Seoul, Korea. In 1998, he earned a Bachelor of Economics in statistics degree from Chung-Ang University, Seoul, Korea. He entered the graduate program in statistics at the University of Florida in August, 1998. He earned a Master of Statistics degree from the University of Florida in August, 2002. In addition to pursuing his Ph.D. in statistics from the University of Florida, he has served as a teaching assistant and research assistant of the Department of Statistics at UF. Also with a motivation, he passed the actuarial sciences exam.

He got married in 1999 and his wife gave birth to a lovely daughter in 2001 during his work. After graduation, the author will begin his next adventure as an actuary at the Samsung Fire and Marine Insurance Co. in Seoul, Korea.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Malay Ghosh

Malay Ghosh, Chair
Distinguished Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ronald H. Randles

Ronald Randles
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Andrew Rosalsky

Andrew Rosalsky
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Cynthia Garvan
Research Assistant Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Beverly Brechner
Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 2004

Dean, Graduate School